Quality Assessment For Hyperspectral Airborne Systems

From Data Up To Land-Products

DISSERTATION
zur Erlangung des akademischen Grades
Doctor Rerum Naturalium
(Dr. rer. nat.)
im Fach Informatik

eingereicht an der
Mathematisch-Naturwissenschaftliche Fakultät
der Humboldt-Universität zu Berlin

von
Ingénieur Généraliste Supaéro,
Diplômé de l'Institut Supérieur de l'Aéronautique et de l'Espace:
Grégoire Henry Gérard Kerr

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftliche Fakultät:
Prof. Dr. Elmar Kulke

Gutarter/in:
1. Prof. Dr. Ralf Reulke
2. Prof. Dr. Thomas Udelhoven
3. Dr. René Preusker

Tag der Verteidigung: 30.07.2015

Short Abstract

This work proposes a methodology for performing a quality assessment on the complete airborne hyperspectral system, thus ranging from data acquisition up to land-product generation. It is compliant with other quality assessment initiatives, such as the European Facility for Airborne Research (EUFAR), the Quality Assessment for Earth observation work-group (QA4EO) and the Joint Committee for Guides in Metrology (JCGM). These are extended into a generic framework allowing for a flexible but reliable quality assessment strategy. Since airborne hyperspectral imagery is usually acquired in several partially overlapping flight-lines, it is proposed to use this information redundancy to retrieve the imagery's internal variability. The underlying method is generic and can be easily introduced in the German Aerospace Center DLR's hyperspectral processing chain. The comparison of two overlapping flight-lines is not straightforward, should it only be because the presence of geo-location errors present in the data. A first step consists in retrieving the relative variability of the pixel's geo-locations, hence providing pairs of pixels imaging the same areas. Subsequently, these pairs of pixels are used to obtain quality indicators accounting for the reproducibility of mapping-products, thus extending the EUFAR's quality layers up to land-products. The third stage of the analysis consists of using these reliability results to improve the mapping-products: it is proposed to maximise the reliability over the mapping-methods' parameters. Finally, the repeatability assessment is back propagated to the hyperspectral data itself. As a result, an estimator of the reflectance variability (including model-, and scene-induced uncertainties) is proposed by means of a blind-deconvolution approach. Altogether, this complements and extends the EUFAR quality layers with estimates of the data and products repeatability while providing confidence intervals as recommended by JCGM and QA4EO.

Keywords— Airborne Remote Sensing, Hyperspectral Processing Chain, Imaging Spectroscopy, Quality Assessment, Signal Processing, EUFAR, QA4EO

Zusammenfassung

Im Rahmen dieser Arbeit wird ein Konzept entwickelt und umgesetzt, welches eine umfassende Bewertung von Daten flugzeuggetragener hyperspektraler Systeme ermöglicht. Es baut auf mehreren aktuellen Initiativen zur Erfassung der Datenqualität flugzeuggetragener Sensoren auf: Der 'European Facility for Airborne Reseach', der 'Quality Assessment for Earth Observation Workgroup' und dem 'Joint Committee for Guides in Metrology'. Bei der Befliegung eines Gebietes mit hyperspektralen Sensorsystemen werden mehrere, teilweise sich überlappende, Flugstreifen aufgenommen. Es wird vorgeschlagen, die Bildinformationen dieser Überlappungsbereiche als redundant anzusehen und so die innere Variabilität der Daten zu erfassen. Die jeweils zwischen zwei Flugstreifen auftretende Variabilität kann (aufgrund unterschiedlicher Blickrichtungen) als ungünstigster anzunehmender Fall ('worst-case') betrachtet werden und ergänzt daher existierende Ansätze, die sich auf die Auswertung homogener Flächen konzentrieren. Das entwickelte Konzept ist auf unterschiedliche Sensorsysteme anwendbar, somit generisch und kann problemlos in die aktuelle Datenprozessierungskette des Deutschen Zentrums für Luft- und Raumfahrt e.V. integriert werden. Im ersten Abschnitt der Arbeit wird dargelegt, wie korrespondierende Pixelpaare, die in den jeweiligen Streifen an gleicher Geolokation liegen, ermittelt werden können. Darauf aufbauend erfolgt eine Plausibilitätsüberprüfung der erfaßten Pixelpaare unter Verwendung von Zuverlässigkeitsmetriken, die auf Basis höherwertigerer Datenprodukte berechnet werden. In einem weiteren Schritt werden die Ergebnisse genutzt, um die notwendigen Parameter für eine optimierte Bildauswertung - hier im Sinne der Zuverlässigkeit - abzuleiten. Abschließend werden die Pixelpaare benutzt, um die globale Variabilität der Reflektanzwerte abzuschätzen. Insgesamt werden durch diese Arbeit die existierenden Methoden zur Qualitätskontrolle optischer Bilddaten umfassend ergänzt.

Schlagwörtern — flugzeugetragene Sensorsysteme, Prozessierungskette für hyperspektrale Bilddaten, spektroskopische Bildaufzeichnung, Qualitätssicherung, Signalverarbeitung, EUFAR, QA4EO

Long Abstract

In order to be able to use results derived from hyperspectral imagery, it is necessary to first validate the corresponding data and products: '*Are they reliable enough for achieving this goal?*', '*What is their quality?*'... Although some initiatives are currently aiming at addressing these questions, several issues remain open. The first one regards the exact definition of 'quality'. In the following the term is restricted to the notion of repeatability. In other words, this corresponds to answering the question: '*What are the discrepancies between two similar measurements?*'

A second issue regards the feasibility of such an assessment. Hyperspectral data allows for the retrieval of a wide range of - possibly subtle - physical processes. Although this is usually an advantage, it also implies that any measurement is influenced by many unwanted sources. As a result, hyperspectral imagery requires a complex pre-processing chain for translating the 'raw' measurements into valuable information. This involves, in turn, further complex considerations. To start with, these pre-treatments are based on models - or approximations of the actual process - they are therefore introducing their own uncertainties into the data, which would in turn also need to be assessed. Besides these limitations, more classical considerations should be made: the sensor itself is not perfect and need to be characterised and calibrated. This stage is also associated with an uncertainty which affects any further step. In order to evaluate the quality of a product - or even to check wether a dataset is suitable for a specific task - the complete system should be investigated: the sensor and the associated characterisation & calibration, the data acquisition, the data pre-processing and the products generation steps. As a result any direct uncertainty propagation within the complete system is - in practice - illusory.

In order to bypass these limitations, the European Facility for Airborne Research (EUFAR) has already produced several sets of quality tags. They do not account for the reliability as such, but only indicate whether such or such aspect of the data is susceptible to affect the quality. Although useful, this information has only been developed for the data itself - up to reflectance: due to the wide variety of products (e.g. classification, quantitative analysis...), mapping applications were left aside. The German Aerospace Center, being part of the EUFAR, is however interested in extending the existing quality tags toward land-products. This work aims at addressing this issue.

Its first steps regards the data quality itself: although of interest, the EUFAR quality tags are not detailed enough for obtaining any quantification of the imagery's reliability. For practical applications, this is theoretically by passed by so-called vicarious (or cross-) validation. The sensor system images a large homogeneous area whose properties are known and compared to the output of the processing chain. However, the huge majority of the areas of interest are heterogeneous, hence bringing further uncertainties (e.g. adjacency effects, influence of differing pixel's footprints). Homogeneous areas are therefore corresponding to an over-optimistic case. Thus, another solution had to be introduced in order to account for more plausible cases.

Since airborne hyperspectral imagery is usually acquired in several partially overlapping flight-lines, it is proposed to use this information redundancy to retrieve the imagery's internal variability. This implies a data-driven quality assessment specific to each data-set. This specificity is however not a limiting factor since this analysis can be performed on virtually every data-set. A further consequence is however the fact that the repeatability assessment is only accounting for relative errors. Again, since the goal is here to investigate the repeatability of measurements, absolute biases are of limited interest - they can actually be assessed by other means. Finally, it is shown that this assessment is actually corresponding to a pessimistic - or worst case scenario - quality assessment, hence complementing existing work on homogeneous areas.

In practice, the comparison of two overlapping flight-lines is however not straight-forward, should it only be because the presence of geo-location errors present in the data. A first step consists therefore in retrieving the relative variability of the pixel's geo-locations, hence providing pairs of pixels imaging the same geo-locations. This stage is performed on atmospherically corrected (reflectance) geo-rectified data. Subsequently, these pairs of pixels are used to obtain 'quality indicators' accounting for the reproducibility of mapping-products. It is in particular advocated to replace the classical Cohen's κ coefficient used in remote sensing by the more generic Krippendorff's α. The latter is indeed more generic: it is not limited to analysis of two classifications - it can for example be used to analyse more than two quantitative mapping results (e.g. organic carbon content of the soil) - while being more stable with respect to configuration changes (e.g. number of investigated pixels). Several other dedicated indicators are also proposed, as well as means to obtain their confidence intervals, hence extending EUFAR results over land-products.

The third stage of this analysis consists in studying whether it is possible to use these reliability results to improve the mapping-products. Once the reliability of a mapping-product has been established, it is logical to to improve it. This work proposes to optimise the reliability over the mapping-methods' parameters. Since this approach is only depending on the reliability retrieved by means of pairs of observations - which can be computed on virtually every data-set - the approach is quite generic.

Finally, the repeatability assessment is back propagated to the hyperspectral data itself. As a result, an estimator of the reflectance variability (including models, as well as scene induced uncertainties) is proposed, hence both complementing the EUFAR quality layers with estimates of the data repeatability and going beyond them by producing quantitative results. This is achieved by means of a blind-deconvolution framework used in conjunction with the use of levy alpha stable distributions.

The proposed approaches are altogether providing generic tools that could be used on almost any data-set with minimal human interaction, hence opening the way to their incorporation within the German Aerospace Center processing chain for airborne hyperspectral data. They are furthermore extensively tested against classical data-sets as well as against a specialised test study elaborated within this work.

Contents

List of Figures

List of Tables

Chapter 1

Introduction & Objectives

> We feel clearly that we are only now beginning to acquire reliable material for welding together the sum total of all that is known into a whole; but, on the other hand, it has become next to impossible for a single mind fully to command more than a small specialized portion of it.
>
> Erwin Schrödinger - Physicist

Chapter Summary

This chapter reviews some considerations about hyperspectral airborne remote sensing. This includes comments and requirements on data and products quality, as well as a clarification of this work objectives.

1.1 Hyperspectral Airborne Remote Sensing

The last decades have seen an increasing use of hyperspectral imagery - also called Imaging Spectroscopy (IS). Applications are now covering a large domain ranging from mineral mapping up to vegetation monitoring or water analysis [Goetz, 2009]. The explanation for this success comes likely from several different reasons. To start with, and similarly to most of the remote sensing sensors, imaging spectroscopy allows for a semi-global coverage of the investigated scene. It therefore produces a comparably homogeneous coverage. Being a 'remote' methodology, it furthermore limits the needs for terrestrial operations, hence greatly facilitating surveys in dangerous or remote areas.

These assets are reinforced by the retrieved amount of data. Imaging spectroscopy is indeed sampling the surface reflectance over hundreds of small wavelength intervals, hence allowing fine estimation of its properties. This can, for example, be used to derive mineral mapping and quantification [Clark, 1999; Teršič et al., 2013b; Žibret et al., 2013] as well as for environment monitoring [Govender et al., 2007; Thenkabail et al., 2000]. Besides this ability for in-depth studies, it can also be used when there is only little prior information. In such cases, IS' high information-content can be used for retrieving which elements are actually observed, hence opening the way to further analysis - including quantitative ones [Rogge et al., 2007].

Although some spaceborne hyperspectral sensors are planned to be launched (e.g. the EnMAP Satellite mission [Stuffler et al., 2007]), there is currently only few fully operational hyperspectral

systems in orbit. As a result, almost every single hyperspectral survey has to make use of airborne sensor systems. The focus of this work is therefore set on airborne spectroscopy, although several concepts could easily be extended to spaceborne systems.

1.2 Objectives

The use of IS in many operational applications - including potentially life threatening ones[1] - calls for an in-depth quality assessment of the system. What is the use of stating "*You are safe*", if one has to add "*But I cannot be sure of it*"? In other words, this corresponds to the liability question from a data - or mapping - producer point of view: *How much can this data be trusted? Am I ready to certify this?*

In practice, such questioning is not specific to hyperspectral systems. The Quality Assessment for Earth Observation work-group (QA4EO) from the Committee on Earth Observation Satellites (CEOS) and the European Facility for Airborne Research (EUFAR) have already issued guidelines and recommendations towards an exhaustive quality assessment. The German Aerospace Center (DLR) being involved in these groups, as well as being a hyperspectral data provider, is naturally interested in developing these concepts further.

The corresponding methodologies have to provide not only an accurate quality assessment, but also to comply with the existing pre-processing chain - as the one developed by DLR. Since the latter has been developed to be comparably automatic and generic, investigations on quality should also be able to be performed in an automatic and generic fashion. They would furthermore have to cover the complete data processing chain: the effects over the complete system have to be investigated: ranging from sensor calibration up to mapping results.

So stated the issue is quite complex. Indeed, since hyperspectral data provides a large amount of information, it is necessarily also influenced by a large amount of uncertainties. The latter are not systematic: they can be scene dependent, hence being hardly addressable in a generic fashion. A further issue concerns the processing chain itself: since it attempts to correct complex processes, it is intrinsically complex too. Although globally improving the data, such models cannot mimic all potential cases and are therefore bringing their own uncertainties [Berk et al., 1998, 2009; Richter and Schläpfer, 2002].

In practice, existing methodologies used for quality assessment are moreover only providing rough quality tags (good/bad) or recommendations without proposing practical methodologies for achieving them (QA4EO). Even these are not covering the complete imaging system. The EUFAR approach is for example only studying the data-quality up to the geocoded reflectance (L2) data. Leaving to one side EUFAR, other approaches have been designed to assess the quality of land-products [Congalton, 2001; Olofsson et al., 2013]: they are however mostly addressing the specific case where ground-truth is available and where the product itself is a classification. These underlying methodological gaps have therefore to be addressed.

This work aims at tackling these issues by starting from the proposed quality tags proposed by EUFAR for the data pre-processing steps. These should firstly be extended toward mapping products. In a second stage, it would be of interest to use these results to improve the mapping results. Finally, gaining an understanding of the data quality beyond 'simple' quality tags is also wished: this would typically lead to providing confidence intervals. Besides being compliant with these initiatives, developments should be applicable to the DLR' hyperspectral processing chain, hence allowing for their incorporation at a later stage. The next section reviews how this was achieved in practice.

[1] E.g. in [Lenart et al., 2011] spectrometry is used in the context of the Kolontár (Hungary) red-mud spill, in the context of emergency-response situations.

1.3 Proposed Approach and Outline of the Work

To start with, the main principles and specificness of hyperspectral remote sensing were reviewed and analysed as done in chapter 2: this includes a survey of the associated processing steps and common tools and practices. Chapter 3 presents then a review of existing initiatives associated with quality assessment. This encompasses not only hyperspectral data, but also other related or more generic methodologies. Based on these results, existing gaps were identified and a suitable bridging methodology was proposed: namely comparing pairs of observations for retrieving a relative quality assessment.

Although promising, this approach requires however that a limitation of the data are treated before being applicable: geo-locations associated with imagery elements are indeed rarely perfect. In practice, large errors might be observed, hence preventing a direct comparison. Chapter 4 presents a methodology for bypassing this limitation: a relative co-registration step is performed and validated. Chapter 5 extends these results by proposing an interpolation scheme of these relative co-registration results over the complete imagery.

Once this initial issue has been addressed, it becomes possible to perform the quality assessment itself. Chapter 6 proposes and reviews several methodologies to assess mapping applications quality, hence extending the EUFAR quality indicators over mapping-products. Details about the corresponding results can be found in appendix B.

Chapter 7 proposes then a methodology for improving the reliability of mapping products based on the methods presented in chapter 6. Finally, chapter 8 introduces an in-depth estimation and analysis of the hyperspectral per-band uncertainty. Chapter 9 summarises at last this work's main outcome, while possible extensions are reviewed in chapter 10.

The link between these steps is outlined in figure 1.1. The notations as well as some definitions used within this work are provided in appendix A while definitions and acronyms are given in appendix H and table H.1. Data-sets and sensor systems used for the investigations are described in appendices F and G respectively. Finally appendices C to E provide an insight of several methods and tools used within this work: statistics and probabilities, Fourier-transforms and linear algebra.

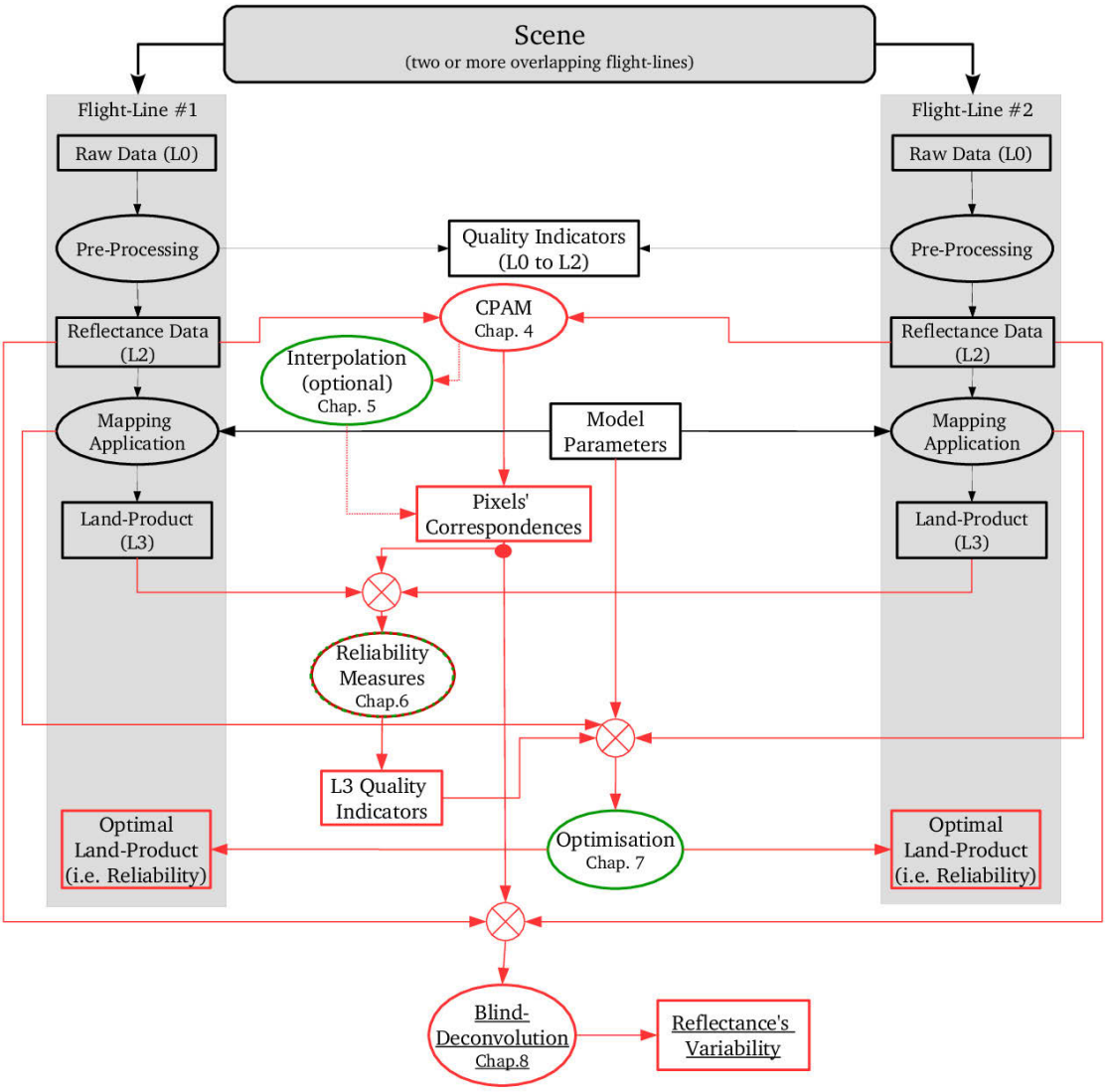

Figure 1.1: Conceptual organisation of this work. Squares indicate data-sets and results, while ellipses stand for the methods. Red elements correspond to the contribution of this work, green ones refer to already existing methods from other fields that were introduced in this work. Black elements represent already established steps. Underlined text stresses the need for a specific acquisition set up.

Chapter 2

Remote Sensing & Spectroscopy

> We have to remember that what we observe is not nature herself, but nature exposed to our method of questioning.
>
> Werner Heisenberg - Physicist

Chapter Summary

This chapter reviews some the notions and methods associated with hyperspectral remote sensing. In a first part generalities about remote sensing are summarised before being further developed in the context of hyperspectral imagery. This is then followed by a review of the hyperspectral processing chain as developed within the German Aerospace Center. A final section presents classical tools used for the generation of land-products.

2.1 Introduction: Remote Sensing

Remote Sensing, should it be airborne or spaceborne, is a convenient Earth observation method [Ustin, 2004]. It does indeed allow for the mapping of large areas with only limited cost and time expense[1]. When a site accessibility is limited, for example due to security or remoteness reasons, it might even be the only Earth observation framework available. Besides a few exceptions, such as mechanical waves analysis (e.g. sonar[2] [Moravec and Elfes, 1985], seismic waves propagation Ben-Menahem and Singh [2000]) or gravimetry (e.g. GRACE satellite mission [Adam, 2002]), the huge majority of remote sensing techniques used for Earth observation are focusing on electromagnetic radiation.

From a strict physical point of view, electromagnetic radiations are to be understood as being made of both a wave and particle (*photon*) nature: this is the so-called wave-particle duality. This distinction is however of little interest for the understanding of this work: unless specified otherwise, electromagnetic radiations are therefore understood as a wave, hence being defined by their wavelength in the vacuum, λ_0. When propagating through a medium the wavelengths are slightly changed by the medium refraction index n: $\lambda = \lambda_0/n$. Since air- and spaceborne

[1] Airborne remote sensing while being actually more expensive than spaceborne, is on the other hand more flexible.

[2] Sonar: SOund Navigation And Ranging

remote sensing usually only deal with propagation in the atmosphere, having a refraction index very close to one, wavelength are considered as being *in practice* constant and equal to their value in the vacuum. An alternate characterisation consists in studying the radiation frequency: $\nu = c_0/\lambda_0$, where c_0 is the speed of light in the vacuum.

Depending on the radiation wavelength several domains are considered (see figure 2.1). In the following, focus is set on the so-called optical domain, ranging from the end of the ultraviolet rays up to the short-wave infra-red (SWIR) where radiation share behaviour of the 'visible light'. Spectroscopy over the optical domain is then defined as being the 'study of light as a function of wavelength that has been emitted, reflected or scattered from a solid, liquid or gas' [Clark, 1999].

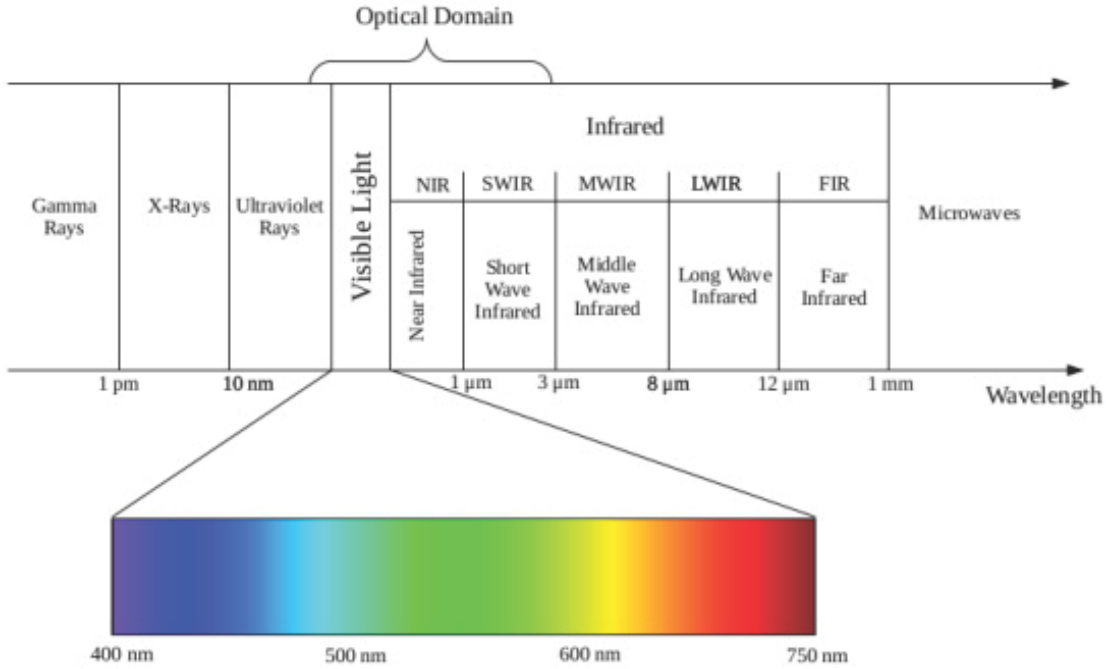

Figure 2.1: Domains of the electromagnetic spectrum and naming conventions.

In the following, focus is set on passive measurement techniques. Such measurements correspond to systems measuring incoming radiations without being their source. In the optical domain, the radiation is typically emitted by the Sun; the sensor measuring their refection on the Earth's surface. Active techniques on the other hand are emitting their 'own' radiation before measuring the returning signal. In the optical domain, this regroups the different LiDAR systems (Light Detection and Ranging).

Section 2.2 is going to summarise the main characteristics of hyperspectral remote sensing, before reviewing the associated pre-processing chain as performed by DLR in section 2.3. Section 2.4 finally reviews some of the most commonly used data analysis tools associated with this kind of measurements.

2.2 Hyperspectral Remote Sensing

2.2.1 Basic Principles

Hyperspectral remote sensing - also called Imaging Spectroscopy (IS) - aims at sampling the electromagnetic radiations over many small portions of the optical and sometimes thermal domains, hence allowing a study of its properties as a function of the wavelength. Multi-spectral sensors are working similarly but do sample fewer, spectrally-broader bands, which may be non-contiguously spaced' [Kramer, 2002].

In airborne and spaceborne hyperspectral remote sensing, the measured radiations are primarily emitted by the Sun before being reflected by the Earth's surface and being collected by the sensor. Although being the main contributor, the Sun is not the only object emitting radiations. According to the Planck's law, an ideal emitter - or black-body - emits radiation as a function of both its temperature and the wavelength: the corresponding radiance - expressed in Watt per steradians per cubic meter - is given by equation 2.2.1:

$$L(\lambda, T) = \frac{2hc_0^2}{\lambda^5}\left(\exp\left(\frac{hc_0}{\lambda k_B T}\right) - 1\right)^{-1} \tag{2.2.1}$$

Where T is the kinetic temperature expressed in Kelvins, $h \approx 6.626\ 10^{-34} J.s$ is the Planck constant, $k_B \approx 1.380 J/K$ is the Boltzmann's constant and the wavelength λ is expressed in meters. Although equation 2.2.1 implies that several other sources might contribute to an observed radiance, the optical domain is mainly influenced by the Sun [Gupta, 2003].

2.2.2 Interactions with Matter

When travelling through matter - including gases - electromagnetic radiations are transmitted, absorbed and reflected, hence splitting the associated energy according to three coefficients of proportionality [Gupta, 2003]: the transmittance τ, the absorbency α and the reflectance ρ. Due to energy conservation principle, these three coefficients are summing to one. They do, however, depend on the material and its state, and on the radiation wavelength.

For the Earth's atmosphere and in the optical domain, the transmittance is predominant. Radiance is however affected by scattering due to aerosols and selective absorption due to atmospheric gases (mostly water-vapour: see figure 2.2).

On the Earth's surface, the dominant effects are on the other hand reflection and absorption. Since both are antagonist, focus is usually set on reflection ρ, the absorption being equal to $1-\rho$. Reflection depends on the surface type and its properties, the radiation's incidence angle and wavelength and are, in the general, case not homogeneously distributed over all directions: such behaviour is described by means of bi-directional reflectance distribution function or BRDF.

Absorbance corresponds to an absorption of the incoming radiation which is typically converted into temperature by the material. In the following, the main object of interest is the Earth's surface reflectance. Radiations measured at the sensor level - *at sensor radiance* - has therefore to be corrected for perturbations such as the radiance alteration made by the atmosphere.

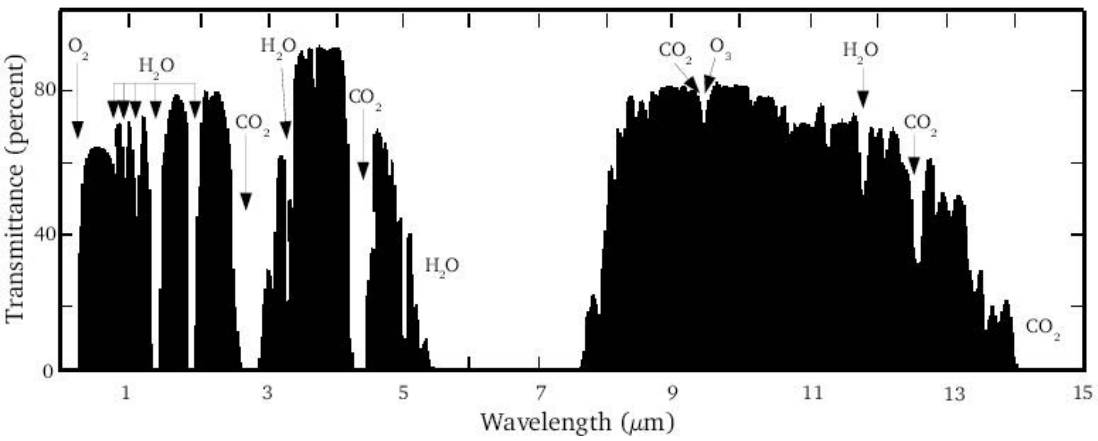

Figure 2.2: Absorptions bands in the optical and thermal domain, in relation with corresponding gases.

2.2.3 Spectral Analysis

The spectral characteristics of the matter - its transmittance, reflectance and absorbance - are ruled by the available energy transitions at atomic and molecular level:

- Electronic process impacting mostly the VIS and NIR domain: the incoming energy brought by the radiation is used to raise the energy level of an electron. The levels are dependant on the electronic configurations of the atoms constituting the material.
- Vibrational process impacting mostly the thermal domain: the incoming radiation is used for 'accelerating' different parts of the molecules consituting the material, hence depending on the material's molecules.

The corresponding reflected spectral features are therefore linked to molecular and atomic configurations of surface materials. Their analysis can therefore be used to gain information about the surface constituents [Bayer et al., 2012; Clark, 1999; Clark et al., 2007; Gupta, 2003]: see figure 2.3.

It should, however, be noticed that spectral features are not only dependant on chemical structures, but also on a wide number of influencing factors such as grain sizes, viewing geometry or mixtures [Clark, 1999; Gupta, 2003; Johnson et al., 1992] as well as physical variability [Messiah, 1999].

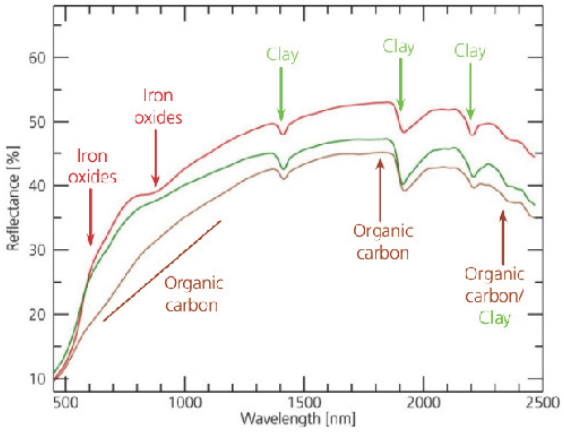

Figure 2.3: Example of influence from soil contents on reflectance: for soils containing mostly organic carbons (brown), clays (green) or iron oxides (red). From [Bayer, 2013].

2.2.4 Measurement Devices

In practice, the complete continuous spectra are actually not measured. Instead, incoming light is split into several small sampling intervals - or spectral bands - and the incoming radiance is integrated over each of them. This integration is typically performed by means of Charged-Coupled Devices (CCD) transcribing radiance into voltage. It should be noted that the energy brought by a single photon is actually proportional to its frequency, ν: $E = h\nu$, with h being the Planck constant. As a result, light at low frequencies (SWIR) conveys less energy, hence being more influenced by noise.

Most airborne sensors - as well as future spaceborne sensors like EnMap [Stuffler et al., 2007] - are whisk- or pushbroom scanning systems [Graham, 2000]. For whiskbroom scanners, spectra are iteratively scanned across the platform track by means of rotating mirror. Spectra are therefore collected pixel by pixel [Fernández-Renau et al., 2005]. When a complete image line has been acquired, the mirror is replaced in its original configuration and the acquisition of a new line is started. On the other hand pushbroom sensors are acquiring the data line by line [Baumgartner et al., 2012]. Due to the platform movement, each line acquisition corresponds to a new area, hence allowing a complete coverage. The principle of a line-by-line acquisition does however imply having a fixed swath. Since it is typically too small to cover the whole area of interest, several flight-lines are typically required to cover the complete site [Habermeyer et al., 2012].

It should be noted that non-imaging spectrometers, acquiring point-wise or continuous profiles of spectra from air or spaceborne platform do also exist (e.g. the Infrared Atmospheric Sounding Interferometer (IASI) [Aires et al., 2002b]). Although the data-analysis is comparably similar [Aires et al., 2002a], the data pre-processing steps as well as the applications are usually differing from hyperspectral imagery. They are therefore left aside. Finally, a few sensor systems are acquiring frame images for each band. Their use has however only been limited so-far [Soukkamaki, 2012].

In conjunction with remotely sensed data, it is not uncommon to perform ground and laboratory measurements: they are usually achieved by means of portable spectrometer measuring punctual spectra [Anderson et al., 2011].

2.3 DLR Processing Chain & Associated Errors

The previous section revealed that although hyperspectral imagery is potentially containing a large amount of information, it is also contaminated by several other influences such as atmospheric ones. As a result, an extensive pre-processing has to be applied. The following sections present the main steps as they are performed within the hyperspectral pre-processing chain developed withing the German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt e.V. - DLR). Other processing chains are however based on very similar principles and involve the same concepts [Beekhuizen et al., 2009b; Bachmann et al., 2011a,b].

It should be noted that this processing chain has been certified with respect to the $ISO-9001-2008$ norm, hence granting a repeatable pre-processing, the underlying references being furthermore calibrated with national and international standards. The different data is moreover provided with suitable metadata, hence insuring a complete traceability [Bachmann et al., 2010b; Habermeyer et al., 2012; Weide et al., 2010] in compliance with European recommendations [Holzwarth and Freer, 2011; European Parliament, Council, 2013].

Level	**Description**
$L0$	Raw data
$L0+$	Raw data, registered to IMU
$L1$	Calibrated, 'at sensor radiance'
$L2-geo$	Geo-rectified, 'at sensor radiance'
$L2-atm$	Non geo-rectified reflectance
$L2$	Geo-rectified reflectance
$L3$	Land product

Table 2.1: Data processing levels and corresponding acronyms (DLR).

2.3.1 Processing Chain Description

The German Aerospace Center is operating a complete hyperspectral processing chain covering the whole system. It ranges from the sensor calibration-characterisation up to land product generation and furthermore includes data acquisition itself as well. The different steps are summarised in the following paragraphs.

Laboratory Calibration

The first step, repeated at regular intervals, is to calibrate the sensor systems. Although called calibration, this is in practice also a characterisation, where the exact characteristics of the sensor are estimated. This is achieved by means of the so-called calibration home base located within the DLR [Baumgartner et al., 2012; Gege et al., 2009; Lenhard, 2012].

The main uncertainties sources are coming from the measurement instruments and standards used. Besides a careful operation and frequent checks, some uncertainties are likely to influence the data [Lenhard, 2012]. A further potential issue is the interpolation: due to the high number of sensor elements to be investigated, a complete characterisation cannot be performed. As a result, these are performed on only a few elements and interpolated over the whole sensor [Gege et al., 2008].

This step is commonly further validated by a vicarious validation where the sensor measurements are compared over well characterised homogeneous surfaces [Secker et al., 2001]. This is

typically achieved by imaging large homogeneous and well known surfaces for 're-scaling' the data on it. Besides this, the sensor characteristics are monitored during the survey, hence allowing the detection of potential issues.

Data Registration

The 'raw' response from the sensor, i.e. a digital signal, is in a second step converted back to radiance, usually assuming a linear relationship. This assumption is comparably right for the mid-range signal, but might lead to slight issues for small and high values. In parallel, each image row is linked with the platform position and attitude, hence opening the way to geometric correction. This is a comparably uncertainty-free processing step [Beekhuizen et al., 2009b,a].

Geo-referencing

Provided the information from the platform lever-arms as well as its attitude and position, it is possible to retrieve from where each pixel was acquired. Given an digital elevation model (DEM) it is then possible to identify the location of each pixel footprint [Müller et al., 2005, 2012; Schläpfer and Richter, 2002; Schläpfer, 2010].

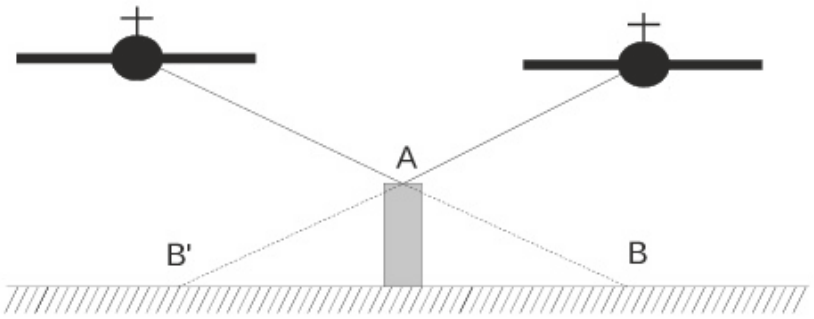

Figure 2.4: Effects of an outdated DEM on ortho-rectification. If the gray-object at A is not taken into-account, the pixels representing it will be geo-localised at B and B', depending on the flight-line.

Under nominal conditions and depending on the system and scenery setup, authors mention that geo-locations accuracies are from less than one pixel's footprint, the airborne platforms being slightly less accurate than spaceborne ones [Müller et al., 2005, 2012; Schläpfer and Richter, 2002]. In practice however the underlying elevation model is not always up-to-date, which might lead to much greater errors [Congalton, 1988; McGwire, 1996; Schläpfer, 2010].

A further source of uncertainty is linked to quick attitude changes of the platform, for example due to wind. If they are of too wide amplitudes, it might lead to gaps in the data coverage, hence leading indirectly in radiometric uncertainties. Further uncertainties sources include the accuracy of the inertial measurement unit (IMU) measuring the aircraft attitude and position. It might for example be improved in presence of differential GPS measurements which are not always available.

Atmospheric Correction

This is by far one of the most difficult parts of the correction chain. It is performed, at DLR, by the AtCor software [Richter et al., 2006]. The goal consists in removing all effects induced by atmospheric transmission. This is typically achieved by iteratively guessing the atmospheric components by selectively studying spectral domain where they are mutually exclusive or well defined as well as making use of radiative models [Berk et al., 2009]. As a result, aerosols' types and content, clouds-cirrus-haze, shadows, visibility and water vapour content can be estimated, provided the imagery incorporates suitable bands [Richter and Schläpfer, 2012].

Once this has been done, effects dependant on the viewpoint and view geometry are removed [Richter, 1997; Richter and Schläpfer, 2002; Richter et al., 2006]. This includes a correction of the reflectance dependency with respect to the illumination and viewing geometry (so-called Bi-directional Reflectance Function or BRDF [Phong, 1975]) as well as signal 'pollution' by non linear light transmission, for example due to scattering.

The atmospheric correction step is extremely dependant on the sensor characterisation step and in particular on the band definitions. Even if these inputs were perfectly error-free, the corrected outputs would contain some uncertainties due to the large amount of models and approximation involved. For example, the use of look up tables of radiative transfer models used in the atmospheric correction is only an approximation [Richter, 1997, 2000], not to mention the accuracy of the radiative models themselves (e.g. 2-5% accuracy within Modtran-4 [Berk et al., 1998, 2009]). The models used for some corrections such as BRDF effects, are moreover comparably simple [Richter, 1997; Richter et al., 2006] with respect to actual BRDF effects [Dorigo, 2008; Peltoniemi et al., 2005]. Finally, most of the correction parameters are linked: uncertainties do therefore not affect one single aspect of the data, but spread over in a non linear fashion.

Ground Measurements

The aforementioned methods are generally associated with ground measurement. Punctual spectra can be collected over areas of interests - or on samples brought to a laboratory - and be compared to the data. Alternatively measurements could be made directly at the level of the mapping-product [Ustin, 2004]. The collection of ground truth is however a complex issue on its own and the corresponding accuracy is not straightforward to estimate [Olofsson et al., 2013, 2014; Foody et al., 2013].

The associated sources of uncertainties regroup localisation errors, sampling schemes for collecting data as well as it representativeness, as well as human error. Besides these issues, the uncertainties of ground-measurements also have to be considered and tackled [Anderson et al., 2011].

2.3.2 Further Sources of Uncertainties

Besides the aforementioned issues, several further technical uncertainties have to be considered. Data is saved digitally: this introduces digitisation noise in addition to the various noise sources of the sensors. The latter can be grouped into the classical noise present in any electrical system (dark current, interferences...) and the ones specific to CCD-imaging systems (shot noise, CCD-readout noise...). Data storage and exchange also introduce potential uncertainties which are extremely hard to assess. They depend on the type of support, their storage environment (e.g. magnetic fields...), the file system used (e.g. journaling or not...) and the file format used (e.g. with or without hash-sum ...). For example, a common file format in the hyperspectral community is the Envi-file-format. In its usual form, it stores the data as a raw binary file without any parity bit check, therefore making it practically impossible to detect unexpected alteration of the data [EXELIS, 2014].

A further source of uncertainty relates to the human factor. This includes manipulation errors as well as software errors in both in-house developed software and external programs. Surveys revealed that only about 8% of scientific software were actually independently validated, not to mention the validation comprehensibility itself [Joppa et al., 2013; Miller, 2006]. This is likely to be even higher for software developed 'in-house' by researchers not necessarily trained in software development standards.

It should be also noted that many issues and errors are not occurring systematically. Digital elevation model errors, might have little or no impact on a scene presenting a flat terrain. On the other hand in a mountainous area, it might have a huge impact. As a result, depending on the scene, uncertainties might have different impact [Beekhuizen et al., 2009a].

Finally most of the uncertainties are actually linked together: for example, an error on the DEM might lead in a false estimation of the adjacency effect [Richter et al., 2006], hence leading to radiometric uncertainties. It results that a manual tracking of all potential errors and influences is technically impossible.

2.4 Common Spectroscopy Tools

Once $L2$ reflectance data has been obtained, the data analysis itself can start. Many tools have been developed for this purpose, of which this section briefly reviews some of the most commonly used ones.

2.4.1 Spectral Metrics

Having spectroscopy data is one thing, making good use of it is another. The first question to arise is then 'how can I state whether two spectra are similar?' - for example when performing a classification. The most straight-forward way to achieve such a task consists in the introduction of so-called 'spectral metrics' suitable for the comparison of spectra among themselves.

This section therefore presents different metrics that could be used to compare the 'distance' between two spectra S and R, each represented by their B sampled reflectance values, assumed to take values in $\mathbb{R}$: $S = \{s_k\}_k$ and $R = \{r_k\}_k$.

It should be noted that some authors prefer not to work directly on reflectance data but to perform a so-called continuum removal [Clark and Roush, 1984]. The convex hull of the spectra are computed and each spectra re-normalised in order to obtain a rectangular function. The main goal is to magnify the absorption features. This process is however hardly justified by physical processes and might lead to poorer results [Filippi and Jensen, 2007].

Note on Norms and Metrics

In its mathematical definition a function $\delta : M \times M \to \mathbb{R}$ on a set M is a metric, or distance function, if and only if it satisfies the following properties for all x, y and z from M:

- Non negativity, or positivity: $\delta(x, y) \geqslant 0$
- Coincidence: $\delta(x, y) = 0 \Leftrightarrow x = y$
- Symmetry: $\delta(x, y) = \delta(y, x)$
- Triangle inequality: $\delta(x, y) \leqslant \delta(x, z) + \delta(z, y)$

If all but the coincidence condition are met, the function is called 'pseudo-metric'. There is a strong link between the notion of metric and the notion of norm. Indeed, a norm N is defined as a function from M to $\mathbb{R}^+$ meeting the following conditions for all x, y and z from M:

- Non negativity, or positivity: $N(x) \geqslant 0$
- Coincidence: $N(X) = 0 \Leftrightarrow x = 0$
- Homogeneity: $\forall \lambda \in \mathbb{C}$, $N(\lambda x) = |\lambda| N(x)$
- Triangle inequality: $N(x + y) \leqslant N(x) + N(y)$

From both properties, it appears that if N is a norm, then $\delta : x, y \to N(x - y)$ is a metric. Indeed, the non-negativity and coincidence properties come directly from the norm properties. The symmetry corresponds to the homogeneity condition taken for $\lambda = -1$. At least the triangular inequalities appear to be equivalent:

$$\delta(x, y) = N(x - y) = N(x + z - z - y) \leqslant N(x - z) + N(z - y) \leqslant \delta(x, z) + \delta(y, z)$$

It should however be noticed that for a metric to induce a norm, some additional properties are required. In particular, if a metric δ verifies for all triplets of points x, y, z:

- $\delta(x, y) = \delta(x + z, y + z)$
- $\forall \lambda \in \mathbb{C}$, $\delta(\lambda x, \lambda y) = |\lambda| \delta(x, y)$

Then $N : x \to \delta(x, 0)$ is a norm. Its non negativity, and coincidence properties are straightforward to check. The homogeneity arises from the second additional requirement. The triangular inequality can be retrieved by combining both additional requirements and the metrics's triangular identity:

$$N(x + y) = \delta(x + y, 0) = \delta(x, -y) \leqslant \delta(x, 0) + \delta(-y, 0) = N(x) + N(y)$$

Classical Norms

In their simplest form, spectra can be seen as vectors from $\mathbb{R}^B$ where N is their number of bands. Such a representation offers a straightforward set of metrics, namely the l^p norms.

$$\forall p \in \mathbb{N}^\star, \|S\|_p = \left(\sum_{k=1}^{B} |s_k|^p \right)^{1/p} \tag{2.4.1}$$

For $p = 2$, one gets the classical l^2 norm, also called euclidean distance, which is commonly used in hyperspectral data analysis.

Technically speaking, since $\mathbb{R}^B$ is a finite dimensional vector-space, all norms are equivalents. In other words, if A and B are two norms, then there exists two strictly positive scalars, α and β such as, for all x in $\mathbb{R}^B$,

$$\alpha A(x) \leqslant B(x) \leqslant \beta A(x)$$

The corollary from this implies that, in spectroscopy, all metrics induced from a norm are equivalent: the metric derived from the l^2 norm is therefore the only one considered in the following chapters.

Spectral Angle

Beside these classical norms, the so-called spectral angle (SA) [Kruse et al., 1993] is defined as the arc-cosine of the classical scalar product of two normalised spectra:

$$SA(s,r) = acos\left(\frac{\sum_{k=1}^{B} s_k \cdot r_k}{\|s\|_2 \|r\|_2}\right) \tag{2.4.2}$$

One can note that the spectral angle is not a real metric. For example, if $r = \lambda s$, with λ being a scalar, the spectral angle between r and s is zero even though both vectors are not identical. The spectral angle is therefore a pseudo-metric.

This mathematical drawback is actually considered as an advantage since changes in illumination can be interpreted as being 'just' a renormalisation of the spectra [Robila, 2004]. From a mathematical point of view, this corresponds to projecting the spectra on the unit length hypersphere centered on zero. The spectral angle can then be understood as the angle defined by the two half lines $[0S)$ and $[0R)$ in the hyperplane including 0, S and R.

It should be furthermore noticed that the spectral angle is not induced by a norm, it is therefore not necessarily equivalent to other metrics such as the ones derived from l^p norms. This does not prevent the spectral angle to be likely the most commonly used spectral metric in the hyperspectral community. It is commonly used as a reference 'metric'. At last, it should be noticed that the spectral angle is not defined for the zero-values spectra.

Spectral Information Divergence

In the previous paragraphs, l^p derived metrics can be seen as a topological understanding of the spectra. On the other hand, the spectral angle could be considered as a geometrical interpretation of the spectral differences. Another interpretation can be based on a probabilistic understanding of the spectra [Chang, 1999].

A spectrum from $[0,1]^B$, can be understood as being a probability vector. The Kullack-Leibler information function could then be used to model the loss of information when approximating a spectrum $S = (s_k)_k$, by a spectrum $R = (r_k)_k$:

$$D_{KL}(S\|R) = \sum_{k=1}^{B} s_k \ln\left(\frac{s_k}{r_k}\right) \tag{2.4.3}$$

A straightforward symmetrisation of the Kullback-Leiber function was introduced as being the spectral information divergence (SID) [Chang, 1999]:

$$\begin{aligned} SID\,(S,R) &= D_{KL}\,(S\|R) + D_{KL}\,(R\|S) \\ &= \sum_{k=1}^{B} s_k \ln\left(\frac{s_k}{r_k}\right) + r_k \ln\left(\frac{r_k}{s_k}\right) \\ &= \ln\left(\prod_{k=1}^{B} s_k^{s_k - r_k} r_k^{r_k - s_k}\right) \end{aligned} \tag{2.4.4}$$

The spectral information divergence can be understood as the sum of the information required additionally from R to obtain S plus the information required to obtain S starting from R. In practice, although the spectral components, s_k are usually already positive (they account for radiance or reflectance), a re-normalisation might be required to obtain spectra from the right domain, namely:

$$S := \frac{S}{\sum_{k=1}^{B} s_k}$$

One could note that, just like the spectral angle, the spectral information divergence is not a real metric either. It is in particular not defined for spectra having bands equals to zero.

In terms of interpretation, and similarly to the spectral angle, the spectral information divergence attempts to get rid of the continuous part of the spectra through the re-normalisation step. It does therefore not discriminate between spectra differing only by a multiplicative factor, hence being less sensitive to illumination changes.

At last, one should note that the spectral information divergence is much more computationally intensive as the other metrics. It indeed requires the evaluation of $4N$ logarithms by call, or one logarithm and $2N$ powers. This therefore limits its use for complex comparisons.

Spectral Correlation

Another interpretation of spectral metrics in terms of information contents was proposed by [De Carvalho and Meneses, 2000]. In its simplest form, the spectral correlation is defined as being the Pearson's correlation coefficient of both spectra. Defining $\bar{s}$ and $\bar{r}$ as the mean values of each spectra, the spectral correlation is defined by equation 2.4.6.

$$\begin{aligned} PC(S,R) &= \frac{\sum_{k=1}^{B} (s_k - \bar{s}) \cdot (r_k - \bar{r})}{\sqrt{\left(\sum_{k=1}^{B} (s_k - \bar{s})^2\right) \cdot \left(\sum_{k=1}^{B} (r_k - \bar{r})^2\right)}} \\ \text{where} \quad \bar{s} &= \frac{1}{N} \sum_{k=1}^{B} s_k \text{ and } \bar{r} = \frac{1}{N} \sum_{k=1}^{B} r_k \end{aligned} \tag{2.4.5}$$

Using the definition from equation 2.4.5, the spectral correlation is absolutely not a metric. In particular $PC(S,S) = 1 \neq 0$! In [Robila, 2004] a 'metrification' of the Pearson's coefficient was proposed, the so-called Spectral correlation Angle (SCA, see equation 2.4.6). This approach does imply that negative correlations imply a greater distance than the absence of correlation.

$$SCA\,(S,R) = acos\left(\frac{PC(S,R)+1}{2}\right) \tag{2.4.6}$$

Just as the spectral angle and the spectral information divergence, SCA is only a pseudo-metric. Two spectra similar up to a multiplicative constant have a distance of zero. Furthermore, and again just SA and SID, the spectral correlation angle is not defined for the zero spectrum.

Total Variation

Under the assumption that the spectra are derivable, one could also use the total variation for discriminating spectra (see equation 2.4.7), which is characterising its gradient variations.

$$\begin{aligned} TVN(S) &= \sum_{k=1}^{B} |\nabla s_k| \\ &\approx \sum_{k=1}^{B-1} |s_{k+1} - s_k| \end{aligned} \tag{2.4.7}$$

One can easily show that the total variation is actually a norm and can therefore be used to define a new spectral metric, namely $TV : S, V \rightarrow TVN(S - R)$. Readers familiar with applied mathematics might also notice that, so stated, the total variation norm is indirectly related to the 'classical' norm on the Sobolev Spaces:

$$\|S\|_{H^1} = \sqrt{\sum_{k=1}^{B} |s_k|^2 + |\nabla s_k|^2} \tag{2.4.8}$$

Closely linked to the total variation, the spectral gradient angle (SGA) was defined by [Angelopoulou et al., 1999] as being the spectral angle between the spectra's gradients:

$$\begin{aligned} SGA(S, R) &= SA(|\nabla S|, |\nabla R|) \\ &\approx acos\left(\frac{\sum_{k=1}^{B-1} |s_{k+1} - s_k| \cdot |r_{r+1} - r_k|}{\sqrt{\left(\sum_{k=1}^{B-1} (s_{k+1} - s_k)^2\right) \cdot \left(\sum_{k=1}^{B-1} (r_{k+1} - r_k)^2\right)}} \right) \end{aligned} \tag{2.4.9}$$

The SGA is de facto ignoring any constant part of the spectra (through derivation) and any multiplicative constant (through the normalisation). On the other hand, total variation based metrics are 'only' removing the constant part of the spectra. It is therefore also a pseudo-metric.

Just like the spectral angle, the spectral gradient angle is actually a pseudo-metric which is not defined over the zero spectrum. On the other hand, $\|\bullet\|_{H^1}$ is indeed a norm, hence producing a 'real' metric. It should be noted that the H^1 norm is intrinsically a norm for functions, therefore on an infinite dimensional space and not equivalent to l_2.

Conclusions on Spectral Metrics

Most metrics defined for comparing spectra are actually pseudo-metrics ignoring the coincidence property. They do indeed ignore multiplicative coefficients (SA, SCA, SID, SGA) or additive constants (SGA). Another interesting feature is that SA, SID and SGA are not defined for the spectrum made of zeros only.

It should be noted that the six presented metrics are actually investigating comparably different properties of the data. SA is merely an angle between two spectra coded as vectors, SCA is directly linked to the correlation of two spectra, SID is based on the information content of the spectra's differences, l_2 is the distance between two spectra coded as points in $\mathbb{R}^B$, and H_1 norm compares the spectra in terms of amplitude and smoothness of their difference. The SGA on the other hand has no direct physical interpretation.

Besides these considerations, some remarks on the computational complexity should also be done. Although all metrics have an overall complexity of $\mathcal{O}(B)$, the corresponding multiplicative

constants differ considerably. For example while l_2 requires only $2B$ multiplications and $2B$ additions (plus a square root evaluation), SID requires $2B$ logarithms, $2B$ multiplications and $2B$ divisions. Table 2.2 summarizes experimental execution times. To be noted is the comparably bad results for the SGA, probably due to the presence of an absolute value.

Metric	Execution Time	Comment
SA	71.1 ms	Fast
SCA	139.1 ms	Medium
SGA	368.4 ms	Slow
SID	5116.4 ms	Extremely Slow
H_1	105.9 ms	Medium
l_2	61.7 ms	Fast

Table 2.2: Spectral metrics complexity: average execution times on 100000 synthetic spectra made of 2000 bands.

2.4.2 Band Ratios

Section 2.4.1 has presented some metrics aimed at comparing spectra one to one. Although using such metrics are necessary for performing in-depth classifications, where each pixel is compared with a reference library of spectra, it is also possible to derive information directly from each spectrum. A common set of tools are the so-called band-ratios. A summary of the most common ones is provided in table 2.3.

The underlying goal of band ratios is to approximate spectral features by linear - or almost linear - relationships between a few spectral bands. Although seemingly crude, they have proved to be comparably reliable tools. For example, the normalised difference vegetation index (NDVI) is commonly used to obtain vegetation masks.

In practice, most of the band ratios are designed to be stable by multiplication. For example, multiplying the reflectance by a constant factor does not change the corresponding NDVI or the ratio vegetation index values (PRI). A few ones, such as the Cellulose Absorption Index (CAI) are stable by adding a constant to the reflectance values. In the latter case, they are usually not stable by multiplicative constants (except for the Leaf Chlorophyll Index - LCI).

Although being simple to use and to port from one sensor to another, it should be stressed that they are not completely stable. They can for example be affected by changes of viewpoint [Hilker et al., 2008].

2.4.3 Conclusion

This review of the most basic tools used for producing $L3$ products from spectroscopy data have revealed that most of them aim for an independence with respect to multiplicative factors, should it be due to a re-normalisation step or due to an explicit division. This can in turn be explained by the presence of BRDF effects: from band to band, the main source of variability is due to the change of viewpoint, which can in turn be modelled as a multiplication of the reflectance.

It should finally be noted that some higher-level tools can be used, especially when quantitative information has to be retrieved [Bayer et al., 2012; Ehrler, 2014; Rogge et al., 2007]. These are, however, usually less generic and are typically specialised toward specific case studies.

Acronym	Name - Explanation	Definition	Reference
CAI	Cellulose Absorption Index	$0.5\,(b_{2000} + b_{2200}) - b_{2100}$	[Nagler et al., 2000]
CSI	Carter Stress Index	$\frac{b_{694}}{b_{760}}$	[Carter and Miller, 1994]
DWSI-5	Disease Water Stress Index	$\frac{b_{800} - b_{550}}{b_{1660} - b_{680}}$	[Apan et al., 2004]
GI	Greenness Index	$\frac{b_{554}}{b_{677}}$	[Zarco-Tejada et al., 2005]
LCI	Leaf Chlorophyll Index	$\frac{b_{850} - b_{710}}{b_{850} - b_{680}}$	[Datt, 1999]
mND_{705}	Chlorophyll Index - 705	$\frac{b_{750} - b_{705}}{b_{750} + b_{705} - 2b_{445}}$	[Sims and Gamon, 2002]
NDVI	Normalised Difference Vegetation Index	$\frac{b_{864} - b_{671}}{b_{864} + b_{671}}$	[Rouse et al., 1973]
NDWI	Normalised Difference Water Index	$\frac{b_{860} - b_{1240}}{b_{860} + b_{1240}}$	[Gao, 1996]
PRI	Photochemical Reflectance Index	$\frac{b_{529} - b_{569}}{b_{529} + b_{569}}$	[Gamon et al., 1992]
REIP	Red-Edge Inflection Point	$700 + 40\frac{0.5(b_{670} + b_{780}) - b_{700}}{b_{740} - b_{700}}$	[Guyot et al., 1988]
RVI	Ratio Vegetation Index	$\frac{b_{864}}{b_{671}}$	[Jordan, 1969]
SR_{705}	Chlorophyll Index	$\frac{b_{750}}{b_{705}}$	[Sims and Gamon, 2002]

Table 2.3: Summary of common band ratios. b_λ indicates the reflectance value of a band located at a wavelength of λ nanometers.

2.5 Conclusions

This chapter has presented an insight into hyperspectral data processing, as performed by the German Aerospace Center. It should, however, be noted that these notions are comparably standard and are used in a similar manner within most of the other hyperspectral processing chains [Beekhuizen et al., 2009b]. Although being certified for its repeatability and making use of traceable references, the processing chain is nevertheless extremely complex and is likely to involve many sources of uncertainties occurring at virtually every single step. To make things worst, uncertainties are linked together and dependant on the correction models themselves.

Chapter 3 is therefore going to review the existing quality assessment methodologies and will put them in relationship with hyperspectral data, therefore allowing for a clearer statement of which issues remain to be addressed.

Chapter 3

Quality Assessment

> My company has had a safety program for 150 years. The program was instituted as a result of a French law requiring an explosives manufacturer to live on the premises with his family.
>
> Crawford Greenwalt - DuPont President

Chapter Summary

This chapter reviews current practices about quality assessment. As a first step, generic definitions are set up. This is followed by two sections reviewing generic quality approaches as well as methods specific to spectroscopy. Finally a conclusion is made by identifying remaining gaps and issues as well as proposing a methodology to bridge them.

3.1 Introduction

Quality assessment has been an extremely widely studied topic and has therefore lead to many different definitions and viewpoints tackling different situations and problems. For example, the Common Assessment framework [Dearing et al., 2006] is aiming at improving public services performances. On the other end of the applications spectrum, the Federal Aviation Agency is defining quality as safety with respect to addressing operational risk issues [AFS, 2004] while the consortium for IT Software quality summarises it as Reliability, Efficiency, Security, and Maintainability [CISQ, 2012]. This chapter aims at clarifying these notions, as well as presenting initiatives suitable for performing quality assessments on hyperspectral remote sensing data.

Section 3.2 defines several concepts related to quality assessment. Section 3.3, in a second stage, reviews generic recommendations for evaluating the quality of a data-set. It is complemented in chapter 3.4 by an review of the non-standard, although widely used, methodologies suitable for evaluating remote sensing data. Finally section 3.5 summarises the remaining gaps and limitations of these methods and propose a complementary framework for a data-driven quality assessment methodology. Appendix H provides the corresponding definitions for the terms which are used in the following sections.

3.2 Generic Concepts

This section presents requirements and constraints related to quality assessment. It makes use of terms and definitions related to quality assessment, whose exact definition is provided in appendix H.

3.2.1 Quality Assessment: Uses and Interest

Before going further, let us consider the case where a temperature, say 37 Celsius, is given. It is in general an interesting information. However different use cases will warrant different levels of tolerance to associated uncertainties:

- An empirical determination of fifths decimal gas-constant according the perfect gas law. The uncertainty should be at the very least smaller than 10^{-5} Celsius (see section A.5.2.
- The effects of global warming over the average temperature in term of climate. For being of significance, its uncertainty should be of much lesser amplitude than the actual changes, which is less than 0.1 Celsius [Tett et al., 1999].
- The human body temperature. The uncertainty should also be comparatively low - say, less than 0.5 Celsius - for allowing to draw useful statement about the presence or lack of fever.
- The monitoring of water temperature in shower: an error or bias of up to one or two Celsius would be fair.
- A stress test of construction materials with respect to fire: up to 10 Celsius of error or bias would not be important.

The aforementioned examples show that the accuracy statement associated with this temperature is mandatory for being able to use the data at all. The following sections reviews this requirement in more details.

3.2.2 Quality for Data Processing

Before stepping into the details of the duality between value and accuracy, one should start by considering an even more generic issue. The German Aerospace Center is to a large extent processing hyperspectral data up to at least $L2$ level. Due to the complexity of this processing chain, it is in turn vital to check whether it produces valid results. How can it be checked whether the produced data is valid, when no standardised quality assessment technique has been made available?

On the longer run, it is furthermore interesting to get to know which parts of the processing chain are contributing to the most uncertainty. This would in turn allow a potential improvement of the results by selectively improving or removing bottlenecks.

3.2.3 Quality for Data Provider

The German Aerospace Center is moreover, to some extent, a data provider. The remote sensing data and eventually the land-products which have been derived from it are not used directly by the institution. They might be provided to other stakeholders for further processing (e.g. as inputs for climatological modelling) or be directly used as input for decision making (e.g. by a government).

If the data is only provided 'as it is' and without any explicit statement of fitness for any purpose, they are useless for the aforementioned users. It would indeed correspond to the statement: "*Here is some information, we do however not know what they are worth*".

On the other hand, under- or overestimating the validity of decision-support data, might lead, in the worst case, to deaths and eventually to the prosecution of stakeholders. An infamous example is the catastrophic Aquila earthquake in 2009 and the subsequent prosecution of several scientists and officials for failing to give sufficient credit to the presaging small tremors [Povoledo and Fountain, 2012][1]. Even if this is definitively an extreme case, this stresses the fact that the quality assessment should be not only performed, but also closely checked and rated. This does in turn discard any manual analysis.

In an ideal world, remote sensing data providers should therefore be able to provide not only the data, but also an assessment of its quality in order to allow decision makers to judge its fitness for purpose. This reasoning is inline with resent activities such as the Global Earth Observation System of Systems (GEOSS) and the quality assessment initiatives from the committee on earth observation satellites (CEOS) [QA4EO task-team, 2010]:

> Data and derived products shall have associated with them an indicator of quality to enable users to assess their suitability for particular applications, i.e., their 'fitness for purpose'.

The German Aerospace Center, being a partner of these initiatives, of course aims at achieving this goal.

3.2.4 Quality as Traceability

Both aforementioned points are leading to a third - almost implicit - one. In order to allow for a comparison, both measurands should be obtained by the same method, up to the few, well defined parameters which have been altered. A corollary is that the processing chain should be always run in the same manner and be based on completely traceable standards.

This can however be achieved relatively easily by setting up a careful process and standards management system. For the specific case of the DLR, this has been achieved by means of a $ISO - 9001 : 2008$ certification. This topic is therefore left aside in further considerations.

3.2.5 Quality as Reproducibility

A final aspect of quality assessment concerns the reproducibility of the measurement. Although spaceborne remote sensing might have the ability to image the same spot at regular intervals, airborne remote sensing usually provides, in the best case, a few surveys each year. This constitutes a further interrogation: 'what if the same survey had been performed one week earlier?'

It should be noted that this is not strictly linked to the intrinsic system-uncertainty - or repeatability. In the latter, only the uncertainty introduced by the system itself is accounted for. In practice, the observed scene is however likely to change by itself over the time. The

[1] This is a press article, and was therefore not peer-reviewed.

atmospheric aerosol content might change from day to day, wind might lay weeds downs. The platform itself might furthermore fly on a slightly different path, hence imaging the scene under a different angle.

This uncertainty should also be accounted for since it also impacts what is actually retrievable from an observation. It has, for example, been shown that vegetation reflectance can change within hours or days due to internal processes [Grace et al., 2007]: if this variation is greater than the observed parameter, trying to measure it makes no sense. The quality of a measurement over smaller period of time should therefore be assessed and would likely differ from the absolute repeatability.

3.2.6 Profane Requirements and Conclusions

Besides these goals, several unrelated requirements should be stressed. To start with, any quality assessment should be feasible: this means that it should be compliant with the existing processing chain. The latter being fully automatic, quality assessment methodologies should also require little or no human interaction. They should finally be generic enough to be applicable to any data-set, hence allowing for a simpler traceability.

As a concluding note, it should be noted that it is in general always possible to improve the quality. However a perfect quality - zero error, zero bias - is never achieved.

The first reason is of a financial nature. Improving the quality is indeed involving a cost in terms of time and efforts. Quality is therefore limited by these constraints. From a physical point of view two more constraints have also to be considered.

The first one relies on the fact that the observation - or measurement - is actually modifying the scene. Without entering into the theoretical details [Schrödinger, 1935], this can be understood as follows. Imagine a ground based spectrometric measurement. An operator has to hold a field-spectrometer near to the point of observation: he, as well as the spectrometer itself, reflects and absorb some of the radiations, hence modifying the input radiations of the observed area. One might of course object that these effects are very small, especially when the field work is performed by trained operators [Anderson et al., 2011; Brogaard and Ólafsdóttir, 1997]. They do nevertheless create an absolute minimal value for the uncertainty.

The second issue is the so-called Heisenberg uncertainty principle and is much more problematic. In physical spectroscopy, it is for example well known that time and energy cannot be known at the same time with an arbitrary precision: (also called Mandelshtam-Tamm inequality) [Messiah, 1999]. As at result the standard deviation of the wavelength from a emitted photon due to decay is infinite, even if most of them are 'close to' their theoretical values. Having an ideal spectral band collecting only photons at a specific wavelength would not solve this issue, hence implying that spectral feature will always have a non-zero uncertainty.

3.3 Generic Quality Assessment

This section starts by reviewing the existing standard methods or initiatives related to quality assessment, starting from the most generic ones down to initiatives specialised to hyperspectral airborne remote sensing. At last, section 3.3.4 reviews some connex initiatives.

3.3.1 Recommendations from JCGM

The main recommendation for uncertainty assessments consist of performing the measurement several times under the same conditions [JCGM, 2008b] - type A analysis. For airborne remote sensing, this would however require performing several times the same flight-survey under the same conditions. Given the costs of a flight survey, it would be unlikely to gain sufficient funding to perform a complete analysis. Furthermore, due to the atmospheric (humidity, aerosols, wind, clouds,...) and the geometric changes (e.g. rotation of the Earth and therefore the sun-illumination changes), it is impossible to grant exactly the same acquisition setup over a long period of time. One would therefore have to move to type B analysis.

Recommendation for type B uncertainty analysis comprise analytical computations and Monte-Carlo error propagation [JCGM, 2008a; JCGM, 2008b; JCGM, 2009; JCGM, 2011].

Analytical propagation is not well suited to the complete processing chain. Indeed, given an uncertainty Δx associated with an input value x_0, the output of a derivable model M might be modelled having an uncertainty Δy such as - formally:

$$\Delta y = \sum_{k=1}^{N} \frac{1}{k!} \frac{\partial^k M}{\partial x^k} (x_0) \, \Delta x^k \tag{3.3.1}$$

Where N is the degree of un-linearity to be taken into account in the Taylor expansion. This Taylor series expansion of the underlying models do however require them to be derivative at least up to degree N which is not necessarily the case. For example in the atmospheric processing step some thresholds are made, hence discarding the validity [Richter and Schläpfer, 2002]. Even when this condition is met, the analytical expression of the $N^t h$ order model's derivative can be extremely cumbersome and lead to impractical equations, especially when x is a multi-component vector. Finally such a framework does not account for the model-induced errors.

Monte-Carlo error propagation proceeds in a slightly similar fashion. The uncertainty associated with each input is modelled and several random samples are propagated through the model, hence leading to an empirical estimation of the outputs' uncertainty. Again, Monte-Carlo error propagation are not accounting for modelling issues and are therefore limited as such. Furthermore, due to the large amount of influencing parameters the computational complexity tends to explode. In the very simplified case where a pixel variability would be influenced by only ten independent parameters, each being fully defined by 10 random samples, the total number of scene pre-processing would reach 10^{10}. In the over-optimistic case were each complete processing could be made in one second, this would mean a computation time of over 300 years. At last, the linkage and exact interactions of all inputs parameters are complex, hence making the inputs generation a very complex scientific task by itself.

Experiments performed on simplified cases were nevertheless performed within the EUFAR project. In order to limit the complexity, only a few parameters were supposed to influence the outputs [Beekhuizen et al., 2009b,a]. Results were strongly dependent on the start hypothesis, hence discarding de-facto the methodology. It should however be noted that Monte-Carlo error propagation does have an interest for characterising the calibration uncertainties of the sensor system [Lenhard, 2012]. In this case, input parameters are comparably independent as well as being completely monitored, hence producing reliable results.

As a conclusion, the methods proposed by GUM seem comparably inapplicable to the case of the complete hyperspectral processing chain. The generic recommendation from the GUM: namely that data should be presented with suitable confidence intervals however remains a valuable concept.

3.3.2 Quality Assessment for Earth Observation (QA4EO)

The quality assessment for Earth observation (QA4EO) from the committee on Earth-observation satellites (CEOS) aims at bringing a systematic use of quality assessment. Also initially focusing on satellite data, the proposed concepts can be directly used by virtually any community [QA4EO, 2010]. The core of the initiative is the so-called QA4EO principle:

> Data and derived products shall have associated with them a fully traceable indicator of their quality.

In its documentation, *QA4EO* focusses on the setup of traceable community agreed reference standards which is in turn closely related to cross calibration as it is described in section 3.4.3. It does moreover brush-up the GUM principle, hence being hindered by the same limitations, the goal of the confidence intervals being nonetheless of interest.

3.3.3 EUFAR Quality Layers

Following the mitigated success of the GUM's guidelines for the hyperspectral processing chain [Beekhuizen et al., 2009a], the EUFAR initiative focussed on providing 'only' quality layers of quality indicators. Quality Indicators (QI) are tags, indicating whether a specific aspect of the data can be fully trusted or not. Quality layers are the set of a specific QI defined for each pixel of a given image (see: figure **??**) [Bachmann et al., 2011a,b].

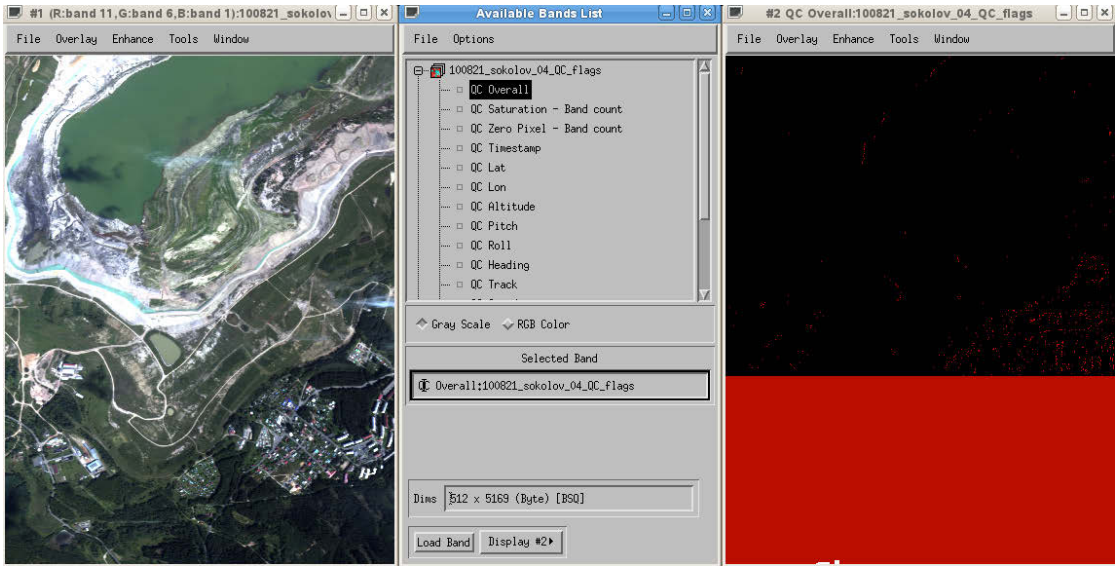

Figure 3.1: Example of a quality layer (right)) representing the 'global quality' QI along with the original imagery (left). Red: there might be an issue, black: no issue detected.

The quality indicators are only providing qualitative information. In their simplest form, they are only boolean information, indicating whether an aspect of the data can be trusted blindly or whether additional investigations should be performed. The EUFAR QI/QL were developed up to $L2$ data: this is, there is currently no quality indicators for $L3$ products. The complete list is provided below.

Sensor Calibration and Data Artefacts

- Aggregated bad pixel mask
- Aggregated interpolated pixel mask
- Saturated pixel / overflow

GPS-IMU Related Errors - Geometry

- Problems with position information and attitude information
- Interpolated position-attitude data

Atmospheric Correction:

- Cloud mask
- Cloud shadow mask (if available)
- Haze mask (if available)

Terrain related:

- Critical local viewing and illumination geometry

Informative layers:

- Water Vapour map
- Aerosol optical thickness (at 550 nm) / Visibility map
- Scan angle file
- DEM (if disposable)

Besides these EUFAR recommendations, DLR has proposed some additional layers:

- Sudden changes in position and attitude
- GPS / DGPS reduced quality
- Land-Water-Shadow mask (apparent reflection or ATCOR)
- Overall quality (Good / Reduced / Low), based on: Saturation, Zero data values, Haze or Clouds presence, Position attitude Issues, Changes in position-attitude, Number of GPS satellites, DGPS availability.

The main advantage of the quality indicators consist in their ability to provide simple information over most of the processing steps. On the other hand, it cannot be used to provide in-depth understanding of the phenomena leading to an issue. As an example, the absence of a DGPS indicates that users might encounter IMU inaccuracies. It does not provide any information on whether or not there was actually a problem. As a last limitation, quality indicators only account for some parameters. Finally, it should be stressed out that quality indicators are not accounting at all for any potential modelling error. For example, the atmospheric QI are directly derived from AtCor outputs, which includes their own uncertainties.

Nevertheless, since these QI have become the agreed standards within the hyperspectral airborne community, they have to be used and is actually extended to the $L3$ products in chapter 6.

3.3.4 Connex Approaches

Due to the complexity of the processing chain, it could be tempting to divide it into several unitary steps with well defined inputs and outputs. Each unit could then be independently tested and validated. Provided some constraints on each of its inputs, it could even be possible to define carefully the uncertainties associated with its outputs. This would actually correspond to a classical software architecture design - or system engineering - as described, for example, in [ECSS, 2013].

The main advantage of this approach consists in its flexibility: changing a specific processing step would allow to just change the corresponding elementary unit, hence removing the need for a complete revaluation of the complete chain. Although interesting, this approach is not really applicable to an already established processing chain: it would involve recreating it from scratch, hence being extremely costly.

3.4 Quality Assessment in Practice

Besides the aforementioned protocols, a number of practical solutions have been developed over the years up to an 'almost-standard' level. This section critically reviews existing methods such as the expert validation 3.4.1, ground-truthing 3.4.2, cross validation 3.4.3, quality indexes 3.4.4, Bayesian methods 3.4.5 and noise estimation 3.4.6.

3.4.1 Expert Validation

A common procedure for validating remote sensing data consists in performing a visual inspection made by one or several experts [Congalton, 2001; Foody, 2010; Foody et al., 2013; Stehman, 2009; Wilkinson, 2005]. This method, besides being time and costs consuming, present the disadvantage of being subject to operators errors and misjudgements. Studies in the fields of aviation and complex systems management indicate that from 70% up to 90% of accidents are due human errors, even in the context of highly trained professionals. This makes humans the main failure contributors by several orders of magnitude [Grabowski and Roberts, 1996; Shappel and Wiegmann, 2000]. In remote sensing, the reported uncertainty varies largely from study to study [Foody, 2010]: [Powell et al., 2004] reported discrepancies of about 30% but with comparably large differences depending on the land-cover type. [Thompson et al., 2007] reported discrepancies ranging from 30 to 60%.

In the field of imagery, this poor performance can be explained by the numerous optical illusions humans are prone to (e.g. see figure 3.2) [Walker, 1973]. It has been shown that human

vision might react differently to the same gray-level colours [Adelson, 1993], or to lines orientations [von der Heydt et al., 1984], depending on the image context. McGurk and Macdonald [1976] furthermore showed that stimuli from different senses might interact with each other and change the perception. Finally, [Segall et al., 1963] indicate that such illusions are likely to be dependant on cultural background, hence making it even harder to assess or correct it.

At last [Eysel, 2003], suggested that images 'in the brain' are usually neither completely fitting to the reality, neither to the perceptions. This brings doubts on the ability of a human operator to perform regular in-depth validation of the data, especially in the context of critical systems and in compliance with traceability requirements. In short, *Errare Humanum est*: error is a human feature and one should not rely purely on human validation for quality assessment.

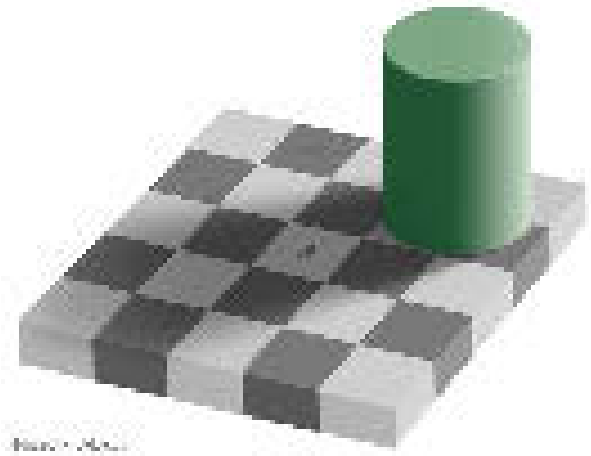

Figure 3.2: Example of a radiometric illusion: squares A and B share exactly the same radiometry although they are perceived differently by humans (©1995, Edward H. Adelson, see [Adelson, 1993] for details).

3.4.2 In-Situ Data

Direct in-situ validation, or 'ground-truth', is also a common way of getting a better understanding of data-quality [Congalton, 2001; Brogaard and Ólafsdóttir, 1997; Foody, 2010; Olofsson et al., 2013, 2014; Stehman, 2009]. The principle consists in checking the air- or spaceborne data versus some measurements performed 'in situ'. This is generally achieved by combining spectrometry measurement acquired directly on the field, ideally during the over-flight, and laboratory measurements performed at a later stage on collected samples. In such a setup, laboratory acquisition, performed in a 'clean' environment are usually considered as a reference, followed by the field measurements. These are then used to calibrated the air- or spaceborne measurements.

In practice, ground spectrometry is not necessary as credible as assumed. Ground spectrometers are also subject to various uncertainties which are not always fully taken into account by operators [Anderson et al., 2011]. When dealing with classification ground-truth, the retrieved ground-truth classes are commonly subject to errors [Foody, 2010; Stehman, 2009] which does in turn affect any quality assessment based on them [Brannstrom and Filippi, 2008; Carlotto, 2009]. Finally non-scientific constraints such as time, costs, or availability might prevent the retrieval of reliable data [Liu and Zhou, 2004]. As a conclusion, [Foody, 2010; Shao and Wu,

2008] reports that ground-truth errors greater than 15% might have to be tolerated. [Johnson and Ross, 2008] reported up to 40% errors due to ground truth geo-location issues even though the ground-truth collecting teams were trained professionals.

At last, the number of ground measurement lies, in practice, between 10 to 100 points and up to a few hundred for extensive studies [Fischer et al., 2014b; Foody et al., 2013; Johnson and Ross, 2008]. These points are usually restricted to homogeneous surfaces. For spectrometric ground truth, the actually imaged surfaces are of less than $10 - 20cm^2$. On the other hand, study areas can commonly be made of about 10^6 pixels, each having a footprint ranging from 0.25 to $16m^2$ and imaging homogeneous as well as heterogeneous areas. This makes a practical extrapolation of the ground-truth data to a complete validation of the imaging data a comparably hazardous task, should it only be from a statistical point of view.

A last argument can be raised on ground-truthing: the whole point of remote sensing consists in performing remote measurements, hence removing the need for an intensive ground-survey. Performing an exhaustive ground-truthing campaign would on the other hand remove this advantage, precisely by performing a ground survey.

Although some vicarious calibration remains definitively required, relying solely on ground-truth seems to be a hazardous task, some authors even preferring to avoid using it [Baraldi et al., 2005; Bruzzone and Marconcini, 2009].

3.4.3 Cross Calibration

The methodologies presented in this section apply primarily to spaceborne sensors. Due to their ability to revisit the same geo-locations at regular intervals, especially over the polar regions[2], it is indeed possible to compare the imagery acquired under similar conditions, and therefore to inter-calibrate them [Chander et al., 2013a,b]. Tools such as the CEOS-COVE have furthermore been developed to help data providers to retrieve data-sets that could be used for cross calibration [Kessler et al., 2013].

Cross Calibration Methods

The basic principle behind cross calibration methods consists in comparing the results obtained by one or more sensors with the results obtained by a reference sensor over the same area. This reference sensor should of course be well characterised and stable over time. For the comparison to be valid, it should moreover be ensured that the two observations are comparable: that is acquired under similar atmospheric, temporal and geometric set-up. As a result, it is usually preferred to perform it over homogeneous areas by a simultaneous nadir overpass, hence imaging the same area at the same time and with the same geometry and allowing to retrieve the bias between the cross calibrated sensors [Cao et al., 2008; Heidinger et al., 2002].

A slightly more advanced methodology consists of accumulating several observations over different times and/or regions and comparing the imagery statistics before rescaling one sensor on the other in order to get homogeneous data. Although this helps producing 'visually' harmonized data, the corresponding gain of accuracy is still under discution [Chander et al., 2013b; Shi and Bates, 2011].

Finally, when simultaneous observations with the same acquisition conditions are not available or feasible, one can rely on the double differencing methods. Both observations are compared by means of an intermediary measurement. In the simplest form, when a sensor A has to be compared with a reference sensor R this can be achieved by introduction a third intermediary B. One could then compare A and B before comparing A and R, hence implicitly leading to a

[2]For the polar orbiting satellites which correspond to the vast majority of Earth observation satellites.

comparison of A and R. This does however require the intermediary sensor to be stable over the comparison time. A simpler alternative can be performed similarly to vicarious calibration by comparing the sensor measurements over a well defined area whose variability can be neglected or at least compensated [Chander et al., 2013b]. The following paragraph reviews such sites.

Cross Calibration Sites

As seen in the previous paragraph, the scene to be used for cross calibration in case of a double difference is to be chosen carefully. In order to allow reliable comparisons, these sites are chosen in order to have high-reflectance as well as being very homogeneous while involving as little BRDF effects as possible. This does in turn imply that the following approaches are limited to a very specific setup, hence not encompassing all the potential variability.

CEOS proposed to use eight test sites - so-called instrumented sites or LANDNET - as standards locations for fields campaign, hence facilitating traceability and inter-comparison [Chander et al., 2013b; Potts et al., 2013]. It should be noted that other test-sites have been used for similar purposes without being in the frame of CEOS [Chander et al., 2007].

Besides these sites equipped for field campaigns, several 'pseudo invariant calibration sites' or PICS usually corresponding to desert sites [Henry et al., 2013; Lacherade et al., 2013] exhibiting high reflectance, no vegetation and low aerosols: they can therefore be used to estimate the stability of an instrument over long periods of time. An alternative to these sites lies in the deep convective clouds (DCC): since they exhibit almost lambertian surfaces with very bright reflectance in the visible domain, they can also be used for cross calibration, a further advantage being their altitude, hence limiting the variability due to atmospheric effects [Doelling et al., 2013]. The latter method is however dependent on the corresponding atmospheric conditions.

Besides this Earth-based sites, several other celestial bodies offer cross calibration potential. For example the Moon presents stable and easy to model reflectance [Kieffer, 1997] which are moreover not influenced by atmospheric effects. Provided that a satellite has enough de-pointing abilities to image the moon, it can therefore valuably be used to cross calibrate them [Stone et al., 2013]. Besides the Moon, several different stars are invariant enough to provide valid calibrations [Dean et al., 2012].

A final class of cross calibration targets consists of variant scenes, but whose variability is very well understood and can be modelled without much errors. For example, over oceans, the visible spectrum corresponding to the top of atmosphere radiance consisted of 90% of Rayleigh scattering due to atmospheric molecules [Vermote et al., 1992]. The latter technique is not performing as well as the PICS or over deep convective clouds but, since the targets are less bright as for the PICS/DCC, it might be of advantage for easily saturated sensors [Chander et al., 2013b]. Alternative methodologies include calibration over liquid water clouds along with a suitable characterisation of their properties. Although the corresponding modelling is slightly more complicated, the wide distribution of theses clouds makes the method comparably advantageous [Ham and Sohn, 2010; Chander et al., 2013b]. Finally the sun-glint method makes use of the higher percentage of solar reflected radiance and can be used for inter-band calibration as well as for MIR band calibration [Chander et al., 2013b; Fougnie et al., 2007; Hagolle et al., 2004].

Conclusion

It should finally be noted that these initiatives are part of an ongoing effort for international cooperation and are by no means a final status [Chander et al., 2013b]. The cross calibration induced uncertainties have been reviewed in [Chander et al., 2013a] for the specific example of Landsat-7 satellite and yielded a total uncertainty of less than 3%. [Thome et al., 2013] performed

a slightly more general investigation and retrieved a similar order of magnitude of about 4.5% uncertainties associated with the process.

Finally it should be noted that pure modelling approaches do also allow to cross calibrate imagery with simulated data [Chander et al., 2013b; Saunders et al., 2013]. This approach is however limited by the simulation models (resolutions, accuracies,...),

The described methods would theoretically apply not only on satellites but also on airborne sensors. For the latter, their ability to image the same area at the same time with several sensors is however comparably limited, the overflight of the PICS and LANDNET sites being usually not feasible due to geographical distance. It is therefore usually preferred to calibrate the data with respect to ground-measurements (i.e. performing a vicarious calibration). It should however be noted that a few satellite-airborne cross calibration have taken place, the costs associated with such a additional flight-campaign on a regular basis been the main drawback of the method [Abel et al., 1993; Newman et al., 2012; Tobin et al., 2006].

As a conclusion, one should finally stress one more time that this class of method is typically addressing a very specific and optimistic case of uncertainty: namely wide homogeneous bright surfaces with little or no BRDF effects.

3.4.4 Quality Indexes

Beside the aforementioned methods, several 'image quality metrics' have been proposed. The goal is here to provide a 'turn-key solution' defining the image quality. It usually comes in form of a mark with an associated scale, each defining whether the data is 'good' or not. A common criticism expressed on quality metric aim at their lack of transparency [Miettinen, 2004]: a comparably 'bad' mark does indeed not necessarily indicate what are the issues. Furthermore, a quality metric only accounts for what it has been designed for. For example, a metric designed on ground sampling distance assessment might be unable to assess effects related to radiometry, although both information are of equal importance.

The best known index is likely the 'National Image Interpretability Rating Scales' or NIIRS which assigns a mark from 0 to 9 to imagery [IRARS, 2012]. The bigger the mark, the more suitable for an object analysis the data. The General Image Quality Equation (GIQE) [Leachtenauer et al., 1997] is an analytical equation allowing to guess the NIIRS score within a standard deviation of 0.3 NIIRS (see: equation 3.4.1).

$$\begin{aligned} NIIRS &= 10.251 - a\log_{10} GSD_{GIQE} + b\log_{10} RER_{GIQE} \\ &\quad -0.656\ H_{GIQE} - 0.344\frac{G_{GIQE}}{SNR_{GIQE}} \end{aligned} \tag{3.4.1}$$

Where GSD_{GIQE}, RER_{GIQE} and SNR_{GIQE} are estimators fo the ground sampling distance, the relative edge response and the signal to noise ratio respectively, H_{GIQE} and G_{GIQE} accounting for edge sharpening side effects. At last, parameters a and b are given as a function of RER_{GIQE} in equation 3.4.2.

$$\begin{cases} a = 3.320 \\ b = 1.559 \end{cases} \text{ if } RER_{GIQE} \geqslant 0.9 \qquad \begin{cases} a = 3.160 \\ b = 2.817 \end{cases} \text{ else} \tag{3.4.2}$$

Besides focusing mostly on imagery geometry, the NIIRS-GIQE has only little plasticity. In particular, the 'edge-sharpening' terms are of no interest in the specific case of hyperspectral remote-sensing where no such processing step takes place. It should finally be noted that the GIQE is not the only predictor of the NIIRS index. For example the Information Theoretic Image Quality Equation (ITIQUE) can estimate NIIRS index based on multi-scale evaluations of the Shannon mutual information [Gerwe et al., 2009].

Besides NIIRS, dozens of quality metrics have been defined, each corresponding to specific data quality aspects. The Phase Quantization Code [Kim and Park, 2010] accounts for image compression and transmissions errors, the image quality metric [Dumic et al., 2010] and Structural SIMilarity index (SSIM) [Wang et al., 2004] for image comparison, the image quality criterion [Jahn and Reulke, 2012] for sensor geometry, [Shnayderman et al., 2006] presented a measure allowing to predict distortions induced by specific noises while [Chen et al., 2008; Piella and Heijmans, 2003; Wei and Blum, 2010] focused on image fusion. Finally Zhang and Luo [2013] proposed an image quality metric based on image comparisons with statistical properties of natural images.

In the domain of airborne hyperspectral remote sensing, [Brook and Ben Dor, 2011] presents two indicators related to the assessment of at sensor radiance: Rad/Ref (At-sensor radiance divided by theoretical ground reflectance) and RRDF (Radiance to Reflectance difference factors), each focusing only on sensor calibration and requiring specific ground truth.

It should at least be noted that 'classical' image metrics such as the signal to noise ratio can also be seen as an image quality metric. Nevertheless, image quality metrics - up to a few exceptions [Jahn and Reulke, 2012; Miettinen, 2004] - are not allowing to link the retrieve criterion with the error sources, and are in all cases focusing on only a few specific applications, hence preventing any generic use.

3.4.5 Bayesian Inference

Bayesian inference and Bayesian networks are based on the so-called Bayes' rule (see equation 3.4.3). It links together:

- the *posterior probability* - $P(H|O)$: the probability that phenomenon, or hypothesis, H, did occur given observation O.
- the *prior probability* - $P(H)$, the probability of the hypothesis in absence of observation.
- the *likelihood* - $P(E|H)$: the probability that the phenomenon leads to the observation.
- the *marginal likelihood* - P(E) the probability of the observation, independently from the phenomenon.

$$P\left(H|E\right) = \frac{P(E|H)P(H)}{P(E)} \tag{3.4.3}$$

Bayesian methods are commonly used in various fields of data-processing. Setting H as being a class and E a pixel value, equation 3.4.3 yields a classification/segmentation methodology (e.g. [Morris et al., 1997]). H being land-surface parameter, e.g. temperature, this yields a parameters estimation framework (e.g. [Morgan, 2005, 2006])...

Although not explicitly stated in the GUM, Bayesian inference can be used for evaluating uncertainty, especially in presence of systematic effects [Lira and Woeger, 2006]. The main advantage of this approach is its ability to bring information 'from other sources' by the mean of the conditional probabilities.

This would however require a complete knowledge of all relationships in the network, which is very unlikely to happen. A further limitation is related to the computational complexity of the task: the Bayesian inference is associated with a NP-complexity, hence making it unsuitable for any high dimensionality system like the hyperspectral processing chain.

3.4.6 Noise Estimation

A last group of methodologies aims at estimating the noise present in the data. Several sub-groups can be identified [Buades et al., 2005].

Homogeneous Surfaces

A straight forward method consists in analysing the imagery over homogeneous areas and considering any variation as being caused by noise, hence corresponding again to a sub-case of vicarious calibration [Delvit et al., 2002; Pagnutti et al., 2002; Scott et al., 1996]. In practice, such large homogeneous areas are not always available. As an alternative to this, Gao [1993]; Reulke and Weichelt [2012] proposed methods for noise estimation on local areas. These methods, besides being more feasible, allow for a noise estimation over the complete radiometric range instead of only focusing on dark and bright targets only.

Noise as De-Correlation

Finally several authors have considered the noise present in spectrometry data as corresponding to the uncorrelated part of the signal variations from bands to bands [Cawse-Nicholson et al., 2013; Corner et al., 2003; Delvit et al., 2002; Sagar et al., 2014]. Similarly, in [Bachmann et al., 2011a], the inter-band correlations are used to detected faulty bands - identified as the de-correlated ones. It should be underlined that - strictly speaking - correlation is usually understood here as being only linear correlation, typically expressed by means of Pearson's coefficient.

Noise as High Frequencies

Finally a last class of authors considers that the noise corresponds to the high-frequencies - or roughness - of the spectra, hence being merely a generalisation of the correlation approach. Spectral-roughness can be filtered by means of low pass filtering [Corner et al., 2003; Savitzky and Golay, 1964] or by decomposition of the signal on a Fourier or wavelet basis where high-frequencies (high details in case of wavelets) are removed [Rasti et al., 2014].

Conclusions

The noise estimation over homogeneous areas is a widely spread methodology whose efficiency is not to be proved. It does however only account for noise occurring over homogeneous areas, hence been unable to account for uncertainties associated with heterogeneous scenes.

High-frequencies and correlation based noise estimations are on the other hand making strong assumptions on the noise and in practice discards the portion of the signal present in this domain. It should be noted that both methods are linked and are the spectral pendant to the noise estimation over homogeneous surfaces: some techniques are actually attempting to merge both approaches [Rasti et al., 2012; Roger and Arnold, 1996].

It should finally be noted that, when a hypothesis is made on the noise shape, it is usually assumed to be Gaussian [Corner et al., 2003; Farzam and Beheshti, 2011; Sagar et al., 2014]. Under such hypothesis, some authors have proposed methodologies to estimate the noise's parameter. For example [Jalobeanu et al., 2002] proposed a maximum likelihood based method to estimate a Gaussian white noise in presence of blur. Some authors have however suggested that the noise might actually be heavy-tailed [Aiazzi et al., 2006].

3.4.7 Miscellaneous Methods

Besides these classical methods, a few other methodologies have to be cited. In [Cerra et al., 2013], the noise over some spectral bands is estimated as being the residual signal once an un-mixing approach provided the amount of signal actually brought by each end-members.

Ground-truth free methods such as DAMA [Baraldi et al., 2005] - as well as similar approaches for the map updating tasks [Bruzzone and Marconcini, 2009] - based on a circular subjective quality assessment have been developed for the case of classification based on remote sensing data. Finally in the field of classification, some authors are bypassing the uncertainty by introducing classifiers taking the data fuzziness into account, the accuracy being modelled by the way the class inclusion strength is partitioned between the classes in relationship with ground-truth data [Foody, 1996; Zhang and Foody, 1998].

3.5 Conclusions and Retained Methodology

From all the techniques presented in the previous sections, two main comments have to be made. To start with, the existing standards have proved to be virtually impossible to use on the complete hyperspectral processing chain. As a workaround, the EUFAR initiative proposed to use 'only' quality tags targeting specific aspect of the pre-processing. Their main limitation lies in the fact that they have not been extended up to land-products. One of the reason for this lies in two subsequent issues:

- How could the large variety of different land products, ranging from pure classification up to materials quantitative mapping, be addressed?
- On what should the comparison be based? For the processing chain, years of expertise have been used to identify the parameters of interests. For mapping applications, each scene has a comparably new setup and would require a new analysis.

Besides this point, one should notice that even though they are not completely standard methodologies, cross calibration methods present a whole range of possibilities. In particular, *QA4EO* is drafting their use as being best practices.

For airborne spectroscopy, access to such calibration sites, or even simultaneous nadir acquisitions are however not possible, at least on a regular basis. A further issue lies in their design: they are typically based on homogeneous areas such as the CEOS Landnet sites. Although this implies a simpler analysis of the data, this does also discard many potential uncertainties that might occur in regular imagery (influence of the actual radiometry [Reulke and Weichelt, 2012], border effects [Richter et al., 2006], ...).

On the other hand, due to the sensor design, hyperspectral airborne data is acquired in several flight-lines. In order to cover the complete area of interest, these are slightly overlapping, hence providing duplicated information over these areas. The following work proposes to use these pairs of observation to provide a data-driven quality assessment for the data.

This assessment is of course only a relative one, hence only corresponding to an 'error analysis' while leaving aside any 'bias assessment'. In practice, this should not be an issue since the biases can easily be retrieved by most of the aforementioned cross calibration methods. On the other hand, since the proposed method should work on the $L2$ data itself, it allows to bypass models, and therefore to account not only for the uncertainties brought by the inputs, but also for the ones eventually brought by the models themselves.

From a more practical perspective, two more points should be stressed. To start with, the proposed methodology is potentially usable on every single data-set presenting an overlap: this

grants a very high generality to the underlying quality assessment. A second point lies in the quality assessment itself. Comparing the data directly to itself has two direct consequences:

- It incorporates not only the absolute uncertainty of the measurement system, but also the uncertainty of the measurand (changes of the scene itself). This is in practice a huge advantage since it links the variability with the scene uncertainty, hence allowing for stating whether a phenomenon can actually be mapped or not.
- It is seen in chapter 4 that this approach actually corresponds to a quality assessment in the worst case scenario, hence complementing vicarious calibration tasks performed on wide, bright homogeneous surfaces, i.e. performing a best case quality assessment.

Chapter 4

Corresponding Pixels Automated Matcher

> A man is about thirty-eight before he stockpiles enough socks to be able to get one matching pair
>
> Merrily Harpur - Writer

Chapter Summary

This chapter reviews how to solve the co-registration issues which might prevent a direct analysis. In particular, an algorithm - CPAM - is proposed for addressing this issue. Right choices for its parameters are then discussed in a second step. A formal validation of the algorithm results is finally proposed before it is illustrated on several examples.

4.1 Introduction

4.1.1 Generic Introduction

As seen in section 3.5, an interesting direction for getting quality assessment consists of assessing the data quality based on data-comparison in overlapping parts of the flight-lines. This theory is based on the assumption that, given a geo-location, it is possible to obtain for each flight-line the pixel that imaged it. Due to the ortho-rectification step, this should not be a problem. In practice, comparably large spatial discrepancies between flight-lines are however still remaining, even when a comprehensive ortho-rectification step took place. For example, in the Sokolov 2010 HyMap data-set, discrepancies of more than 10 pixels can frequently be observed (see figure 4.1).

Reasons for these discrepancies might of course depend on several factors such as IMU errors or lever-arms mis-measurements. The main contributor is however usually the digital elevation model used for ortho-rectification. It might indeed be outdated, or with a much coarser resolution than the imagery, hence preventing a good geo-rectification from taking place. The differences between the elevation model - accounting for the ground surface only - and the actual surfaces elevation - including tree crowns, buildings, etc - might furthermore lead to further discrepancies [Schläpfer and Richter, 2002]. Such discrepancies are therefore not easily corrected without bringing in at least a new surface model, which in turn might be not feasible due to both time

and cost constraints. For archived data-sets, this is simply impossible. One would therefore have to live with these geometric discrepancies.

They are, on the other hand, preventing an in-depth analysis of the data. For example, comparing a pixel representing grassland with a pixel corresponding to concrete would *hide* the actual variability within their spectral differences. In practice, this discards any possibility of further analysis. A better solution would therefore have to be developed. In other words, a methodology ensuring that pairs of pixels brought to comparison are indeed imaging *roughly* the same area has to be used: this implies performing a co-registration before analysing the data.

This co-registration has to be automated and generic, therefore making it usable for the processing of any data-set. The focus is here more to obtain a pixel-to-pixel registration - this is, pairs of pixels corresponding to the same footprint - than image to image registration where the overall images are fitted together. As long as enough pairs of pixels are retrieved to be statistically meaningful, a complete co-registration of the images is not required.

It should however be kept in mind that a perfect overlap of pixels' footprints from two flight-lines is quite unlikely. This is however due to the nature of the raster data and should therefore be accounted as a contribution to the final uncertainty.

Section 4.1.2 reviews existing methodologies, sections 4.2 and 4.3 presents the selected algorithm and its parameters. Section 4.4 then proposes a formal validation of the method. Finally section 4.5 presents some case studies.

Figure 4.1: Example of ortho-rectification issues over the Sokolov test site. Please also note the BRDF effects affecting radiometry between flight-lines

4.1.2 Co-Registration: State of the Art

The co-registration of pairs of images has been a widely studied topic and has therefore lead to several different methodologies [Brown, 1992]. This section presents an overview of the main frameworks along with an analysis of their suitability in this context.

Basic Principles

The vast majority of co-registration methods involves the four following steps [Zitová and Flusser, 2003].

1. Feature Detection: for each image, distinctive features are retrieved
2. Feature Matching: features from both images are compared and matched.
3. Transform Model Estimation: the 'projection' required to project both images on the same reference is computed from the pairs of features.
4. Image Transformation. Images are re-projected and re-sampled using the transformation from step 3.

In this section, steps 3 and 4 are left aside: re-projecting imagery, along with the underlying re-sampling likely introduces other uncertainties in the data. Focus is therefore set on only retrieving pairs of features that could be related to pixels positions.

Zitová and Flusser [2003] consider two main families of methods applicable for performing steps 1: the '*area-based*' and the '*feature-based*'.

Feature Based Methods

Feature-based methods are basing themselves on identifying salient 'objects' in imagery. Such objects can be:

- Large homogeneous areas such as water bodies [Goshtasby and Stockman, 1985] or forests [Sester et al., 1998]. This does however usually require an additional segmentation task for extracting these areas from the input imagery.
- Linear structures such as coast lines [Maître and Wu, 1987] or roads [Li et al., 1992]. The lines can be extracted using standard edge detectors such as the Canny or Sobel filters and/or the Hough transform [Russ, 2002].
- Points or pixel elements. These can be based on local image extrema, corners and curves' inflections [Stockman et al., 1982], time-frequency features [Zheng and Chellappa, 1992], etc. Modern algorithms are usually taking advantage of several of these properties, hence granting a higher discrimination along with a comparable invariance with respect to illumination, rotation or scale changes. Efficient methods are the 'Scale invariant Feature Transform' (SIFT) [Lowe, 2004] and 'Speeded Up Robust Features' (SURF) [Bay et al., 2008].

When keeping in mind the generality requirement, the homogeneous areas and linear structure features can directly be left aside. Such elements are indeed not necessarily present in each scene. It is for example not possible to perform a valid extraction of large water bodies over an arid scene. Point features, and especially SURF and SIFT are, on the other hand, functioning on virtually every not completely homogeneous scene.

Despite being comparably robust to scale, illumination and rotation changes as well as being fast to compute, these feature descriptors have been designed to only return the 'best' features. They therefore only return a comparably small number of highly valuable points. For example, for a scene made of about 1.5 billions pixels, the SURF feature descriptors retrieved only 4508 feature points, this is less than 0.3% (see figure 4.2).

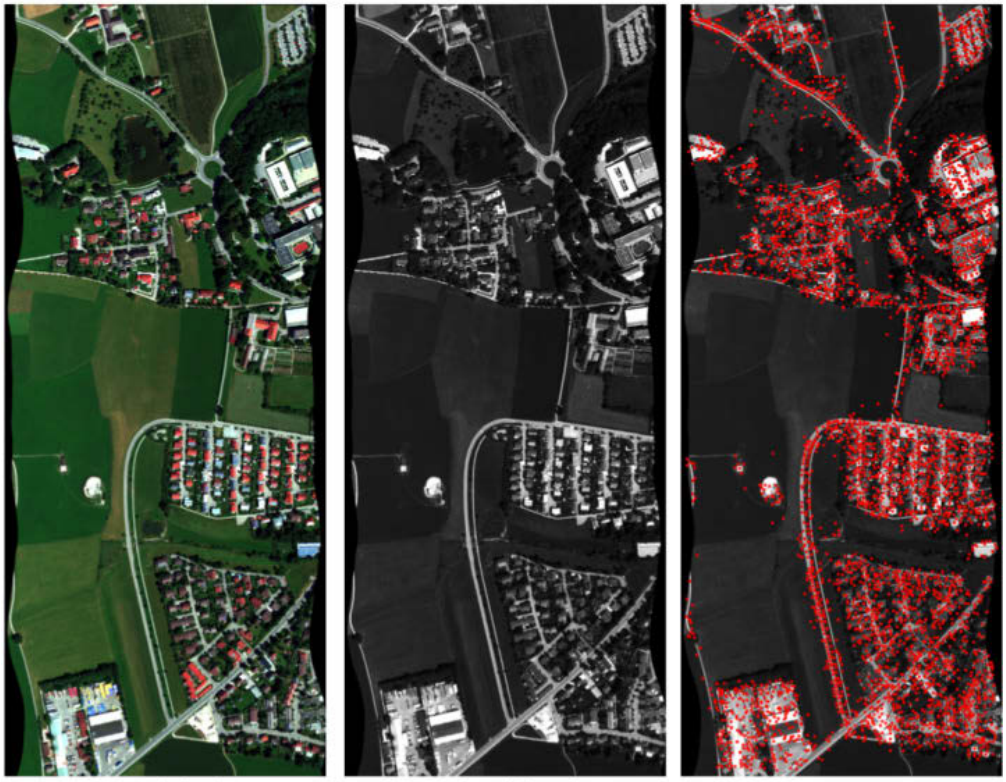

Figure 4.2: Retrieved feature points using SURF (single band) over a Kaufbeuren scene subset (see F.1); 1.5 billions pixels, 4508 feature points retrieved by SURF (red squares).

A further issue lies in the design of these feature descriptors. They have mostly been designed to work on single channel images. Abdel-Hakim and Farag [2006] have proposed an extension of the SIFT to colour images: CSIFT. It is however based on the colour invariance as defined by Geusebroek et al. [2001], which is restricted on 'classical' RGB images and not directly extensible for wavelength outside the visible domain. Fan et al. [2009] have similarly proposed an extension of SURF to colour images. It however makes use of the YUV colour space which also restricts its application to classical RGB colour images.

The two last points are not intrinsic limitations. It should be possible to develop a feature descriptor which is returning a larger number of feature points while still allowing a comparably safe matching. It should all the same be possible to create a feature descriptor actually working not only on RGB images, but also over the complete spectral content. Combining both tasks however seems comparably hazardous, and the area based methods were preferred.

Area Based Methods

Area based methods are merely merging the feature detection and feature matching steps. This is usually done by searching the closest match of each pixel from the first image into the second image.

The classical method consists of defining the goodness of a match as being the maximum of the local normalised cross-correlation (NCC) [Zitová and Flusser, 2003]. Given a neighbourhood size, N, two grey-scale images whose pixel values at column x and row y are respectively given by $I(x,y)$ and $J(x,y)$, the NCC between pixels located at column x_1 and row y_1 in the first image (resp. column x_2 and row y_2 in the second image) is given by equation 4.1.1. Note that it corresponds to the case were translation is the only transformation to be investigated. Additional transformations such as rotation might be brought into the model. The resulting computational cost increase is however in practice preventing this.

$$NCC_N(x_1,y_1,x_2,y_2) = \sum_{k=-N}^{N}\sum_{l=-N}^{N}\frac{\left[I(x_1+k,y_1+l)-\langle I\rangle_{N,x_1,y_1}\right]\left[J(x_2+k,y_2+l)-\langle J\rangle_{N,x_2,y_2}\right]}{\sigma_{I,N}(x_1,y_1)\,\sigma_{J,N}(x_2,y_2)} \quad (4.1.1)$$

where $\langle I\rangle_{x,y}$ (resp. $\sigma_I(x,y)$) represents the mean (resp. the standard deviation) of image I around column x and row y on a square window of width $2N+1$. For each pixel from the first image located at (x_1,y_1), a pixel from the second image is associated by maximising equation 4.1.1:

$$(x_2,y_2) = \arg\max_{v,w} NCC_N(x_1,y_1,v,w) \quad (4.1.2)$$

The normalised cross-correlation, although being the widely used framework, is not the only discrepancy measure available. For example, Viola and Wells [1995] or Pluim et al. [2003] proposed to take advantage of the mutual information (MI) between both images. This can be seen as the spatial pendant to SID - which focuses itself on radiometry. The search for the best match then corresponds to finding the pixel maximising the mutual information. In order to compare subsets from two images, the histograms (p_I and p_J) and joint histogram (p_{IJ}) of their grey levels have to be computed: they are used as approximations of their probability distributions. The mutual information is then defined similarly to the Kullback-Leiber distance between their joint distribution p_{IJ} and their joint distribution in case of complete independence of both subsets: $p_I \times p_J$:

$$MI(x_1,y_1,x_2,y_2) = \sum_k\sum_l p_{IJ}(k,l)\log\frac{p_{IJ}(k,l)}{p_I(k)\,p_J(l)} \quad (4.1.3)$$

The main advantage of this family of techniques is that they virtually allow a matching of every single pixel. On the other hand, due to the search around a given 'neighbourhood', such methods are also more likely to be sensitive to local image distortions like rotation or warping. Furthermore, the greatest drawback associated with this kind of area based method is their very high computational costs. Equation 4.1.2 has indeed to be solved for every pixel from the first flight-line, p, hence leading to a complexity of at least $\mathcal{O}(pN^2)$. This can even increase when complex methods as the mutual information are used and/or when advanced optimisation methods are required to solve equation 4.1.2.

An alternative, especially efficient for large images, consist of taking advantage of the Fourier-transform properties. Consider an image to be a function of two variables - $I : x,y \rightarrow I(x,y)$

- associating the grey level value $I(x, y)$ to the pixel located a column x and row y. It is then possible to apply a two dimensional Fourier transform to the corresponding function I. When having two identical images, just differing by a translation of amplitude (x_0, y_0), I and J, their cross-power spectrum[1] is directly related to the translation by relation 4.1.4. It is then straight forward to retrieve the translation parameters by applying the inverse Fourier transform to the result which yields a Dirac located at (x_0, y_0).

$$\frac{\hat{I}(\omega_1, \omega_2)\,\bar{\hat{J}}(\omega_1, \omega_2)}{|\hat{I}(\omega_1, \omega_2)\,\bar{\hat{J}}(\omega_1, \omega_2)|} = e^{2i\pi(\omega_1 x_0 + \omega_2 y_0)} \tag{4.1.4}$$

Besides being very efficient on large data-sets, the method is very stable with respect to change of illumination. On the other hand, it only accounts for a translational co-registration. De Castro and Morandi [1987] proposed an extension to allow a retrieval of rotational changes. Chen et al. [1994] and Reddy and Chatterji [1996] have extended the methodology to the case of scale changes by means of the Mellin transform.

Due to the uncertainty principle, Fourier transform based methods are however only valid for performing a global co-registration. When co-registration issues are coming from DEM issues, they are quite unlikely to be global and vary through the image. Fourier based methods are therefore of little use in this context: correlation-like methods are therefore preferred.

They have however been mostly defined for single channel imagery. The first step therefore consists of adapting them to hyperspectral imagery, where information is mostly lying in the spectral content of each pixel. Section 4.2 presents the retained approach.

4.2 Proposed Methodology: CPAM

In the following, and for the sake of simplicity, pixels from the first flight-line are noted without a tilde, while their counterparts from the second flight-line are marked with a tilde.

4.2.1 Principle

The point is, in the following sections, to solve the following issue: *Given a pixel from one flight-line, how can the pixel representing the same spot in another flight-line be identified?*.

Since it is here dealt with hyperspectral remote sensing, one can argue that each spot of the ground is defined in a unique way by its spectrum. It should therefore be possible to compare the spectrum from each investigated pixel to the corresponding ones from the second flight-line and only retain the most similar one.

In practice, complexity associated with this straightforward method is extremely high. For a pair of images, each having about C columns and R rows, this yields CR comparisons for dealing with each pixel from the first flight-line. This yields a total complexity of $\mathcal{O}\left(C^2R^2\right)$. For a comparably common HySpex data-set made of 1600 rows and 7000 columns, this implies a total of $1.25\ 10^{14}$ comparisons. Assuming each comparision is made in about one microsecond (see table 2.2), this corresponds to an execution time of about four years.

Another potential issue associated with this approach is its potential for matching together two pixels having very similar spectral content but lying on two very different geo-locations.

In classical stereo-vision, these limitations are usually raised by performing the search along epipolar lines. In other words, given a specific pixel from an image, it is known before hand that the observed point lies 'somewhere' on the line starting in the camera centre and aiming to the

[1] The word '*spectrum*' is here to be understood in its 'Fourier transform' sense.

pixel: the epipolar line. Provided some information about the imaging system geometry and its consecutive positions, it is possible to project this line onto any second image [Hartley and Zisserman, 2004]. If the observed object was actually observed in the second image, it has to lay on the line's projection. When trying to identify it on the second image, one can therefore 'just' search for it along its epipolar line in the second image.

In our case, such detailed information about the system geometry is not available. However, due to the ortho-rectification step, an approximation of each pixel's geo-location is known. And this for each flight-line. Although the ortho-rectification is not perfect, it is *roughly* the right one. In other words the actual geo-location of the pixel lies within a comparably small distance from its theoretical localisation. This property is used to restrict the number of potential matches to be investigated. Given a pixel from the first flight-line, p, and a spectral metric $\mathcal{S}(\bullet, \bullet)$, the search for its 'best' match can be done by means of algorithm 1.

Algorithm 1: Description of the search of best matches: single pass scheme.

Data: A pixel: p
Result: $\tilde{q}$: p's best match

Initialisation:
Compute its theoretical geo-location: l
Compute the pixel from the second flight-line corresponding to geo-location l: $\tilde{p}$
`// We know that the actual match of` p `has to be in vicinity of` $\tilde{p}$
begin
  `// Initialise the 'best match' of` p`:` $\tilde{q}$
  $\tilde{q} := \tilde{p}$
  `// Search for pixels` *`not too far`* `from` $\tilde{p}$
  for $\tilde{r}$ *close to* $\tilde{p}$ **do**
    if $\mathcal{S}(p, \tilde{r}) < \mathcal{S}(p, \tilde{q})$ **then**
      $\tilde{q} := \tilde{r}$
    end
  end
end

The definition of '*not too far*' is clearly referring to a geometric distance. Its exact value is however dependent on the accuracy of the ortho-rectification. For a scene with an ortho-rectification having less than one pixel error, this distance could be one pixel. For scenes presenting deep perturbations, it could be much more. In the following, this value is left as a parameter of the algorithm. Its value is noted w. For the sake of simplicity, the search is not performed on the circle centred on $\tilde{p}$ and of radius w. It is instead performed over the square window centred on $\tilde{p}$ and of width $2w + 1$.

In order to further improve the reliability of the method, yet another extension was performed. When just comparing the spectra pixel-to-pixel, a comparably high risk of mismatch is still present. Due to the presence of noise in the data, locally homogeneous areas could lead to a comparably high amount of mismatches. This can be avoided by comparing groups-of-pixels to groups-of-pixels. The retained solution was to perform computations over a square neighbourhood. Denoting $p_{x,y}$ (resp. $\tilde{p}_{x,y}$) the pixel located at column x and row y in the first flight-line (resp. the second flight-line), the 'spectral metric' over a neighbouring of size n hence becomes

$$\mathcal{M}_{\mathcal{S},n}\left(p_{w,x},\tilde{p}_{y,z}\right)=\sum_{k=-n}^{n}\sum_{l=-n}^{n}\mathcal{S}\left(p_{w+k,x+l},\tilde{p}_{y+k,z+l}\right) \tag{4.2.1}$$

In an ideal world, the choice of n would be comparably easy: the greater, the better. In practice, things are not that simple. First of all, performing computations over a given neighbourhood implies an increase of spectra comparisons by a factor n^2.

Secondly, the computations over the neighbourhood imply that the imagery shares locally the same geometry on both flight-lines, i.e. that a 'horizontal squared neighbourhood' in the first flight-line indeed corresponds to a 'horizontal squared neighbourhood' in the second flight-line. In other words, this implies that the images' relative rotations and deformations are negligible between both flight-lines.

For small neighbourhood sizes this approximation is likely to be valid. Imagery orthorectification grants that there is no huge residual distortion. If any still occurs, it is likely to be hidden in the rasterisation as illustrated in figure 4.3, green and red objects. On the other hand, when dealing with wider neighbourhood sizes, even a small uncorrected distortion might lead to different geometry: see for example the blue object in figure 4.3 in case of an uncorrected rotation. In practice, such distortions might be present, for example due to the issues of the DEM errors and should therefore be considered.

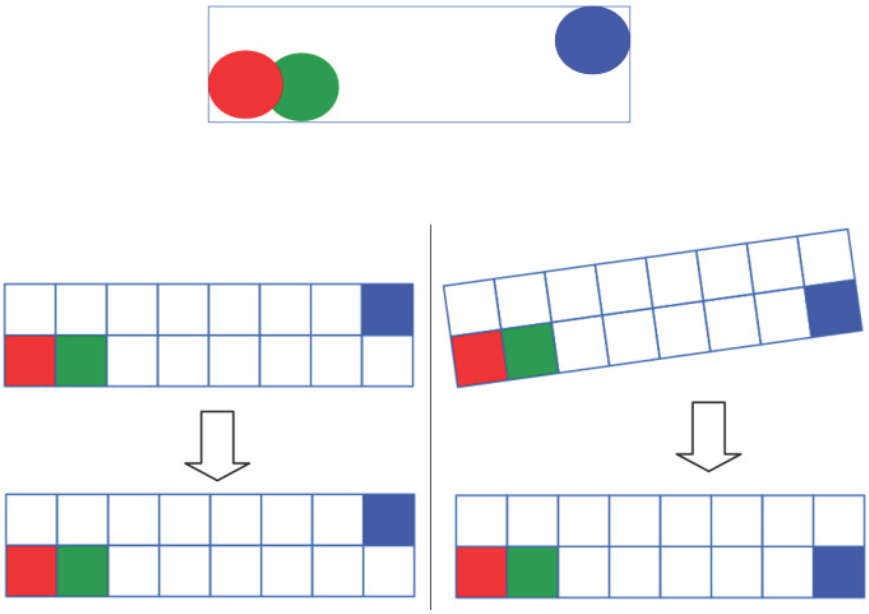

Figure 4.3: Effects of rotations when imaging the scene (top) with different residual errors (left and right). Computing $\mathcal{M}_{\mathcal{S},n}$ on the red pixel over a small neighbourhood will produce better results than computing it over a broad neighbourhood that would include the blue pixel.

A solution to the last issue could of course have consisted in bringing some distortion terms in equation 4.2.1. The search for the best match could then be performed for each possible quantified distortion, hence producing a priori finer results. Such a workaround is however comparably expensive in terms of computations. For example, in order to account only for ten

possible rotations, the execution time is multiplied by ten. In the following, the focus is therefore set on optimising n instead of trying to account for all possible set of distortions.

4.2.2 Validating Found Pixels

Section 4.2.1 has presented a way to compute the best match of a given pixel. The question that now arises is: *How can we be sure that the retrieved pairs do indeed correspond to two observations of the same geo-location?*

Although this task could be done by manual inspection of the results, it remains difficult to perform an automatic and in-depth statement in the general case. Ortho-rectification issues might indeed be location dependant. A direct outlier removal through threshold is therefore out of question and more complex approaches have to be used.

The proposed solution consists of performing a dual search. Indeed, if one is able, for any pixel p from the first flight-line, to retrieve its 'best match' from the second flight-line, $\tilde{q}$, one should also be able to retrieve the best match of $\tilde{q}$ in the first flight-line by using the same approach: call it r.

If r and p are two different pixels, one can safely state that there was at least one issue, no matter whether it was when getting $\tilde{q}$ from p or r from $\tilde{q}$. In such a case, both pairs $(p, \tilde{q})$ and $(r, \tilde{q})$ are marked as unreliable and discarded from further computations.

On the other hand, if p and r are indeed the same pixel, two options arise:

- $p = r$ and $\tilde{q}$ are indeed two observations of the same geo-location.
- $p = r$ and $\tilde{q}$ are not two observations of the same geo-location: this means not only that there was an error when getting $\tilde{q}$ from p and another error when getting r from $\tilde{q}$ but also that these two errors acted as complete opposite.

Since the second case is unlikely to happen, it is neglected. In the event where $p = r$, the pair $(p, \tilde{q})$ is therefore considered as being a reliable pair of corresponding pixels and used in further computations. This approximation might lead to the introduction of a few outliers in the obtained pairs but this effect is ignored for now. The corresponding algorithm for the search for corresponding pixels is given by algorithm 2 while schematic illustrations are provided on figure 4.4 and 4.5.

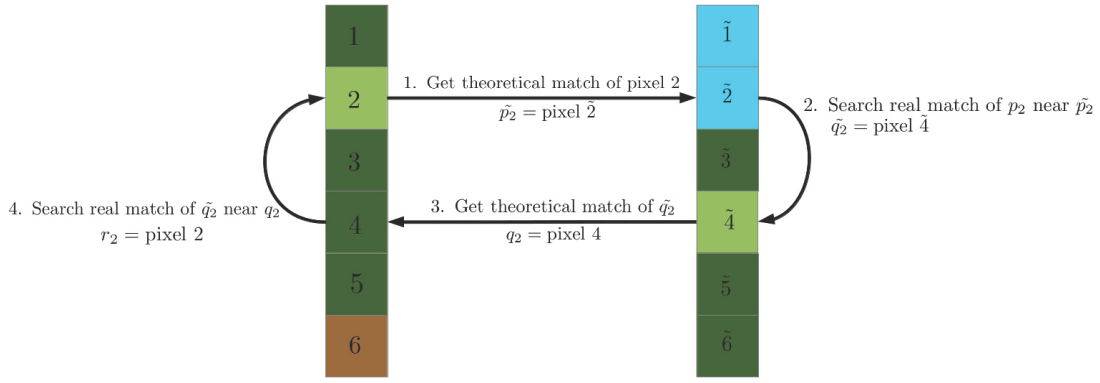

Figure 4.4: Schematic illustration of the correspondence search in the event of a successful retrieval of the corresponding pixels: r_2 is actually p_2, the pair $(p_2, \tilde{q}_2)$ is therefore kept.

Algorithm 2: CPAM algorithm

Data: A spectral metric: $\mathcal{S}(\bullet, \bullet)$
Data: A neighbourhood distance: n
Data: A search window size: w
Result: L the list of retrieved pairs

Initialisation:
$L :=$ empty list
for *Each pixel from the first flight-line:* $\{p_k\}_k$ **do**
  Compute its theoretical geo-location: l_k
  Get $\tilde{p}_k$, the pixel from the second flight-line theoretically located at l_k
  `// Initialise the 'best match' of` p_k`:` $\tilde{q}_k$
  $\tilde{q}_k := \tilde{p}_k$
  for *Each pixel* $\tilde{t}$ *located in the square centered on* $\tilde{p}_k$ *and of width* $2w+1$ **do**
    if $\mathcal{M}_{\mathcal{S},n}(p_k, \tilde{t}) < \mathcal{M}_{\mathcal{S},n}(p_k, \tilde{q}_k)$ **then**
      $\tilde{q}_k := \tilde{t}$
    end
  end
  Compute $\tilde{l}_k$: the theoretical geo-location of $\tilde{q}_k$
  $r_k := q_k$
  for *Each pixel* t *located in the square centered on* q_k *and of width* $2w+1$ **do**
    if $\mathcal{M}_{\mathcal{S},n}(t, \tilde{q}_k) < \mathcal{M}_{\mathcal{S},n}(r_k, \tilde{q}_k)$ **then**
      $r_k := t$
    end
  end
  if $r_k = p_k$ **then**
    insert $(p_k, \tilde{q}_k)$ into L
  end
end

One can easily notice that algorithm 2 is only producing a partial coverage of the imagery. Due to the discarded pairs there are some pixels having no matches. Such pixels are more likely to be found in homogeneous areas where the lack of contrast is intensifying the effects spectral variability. As a result, homogeneous areas are likely to be under-represented in the retrieved

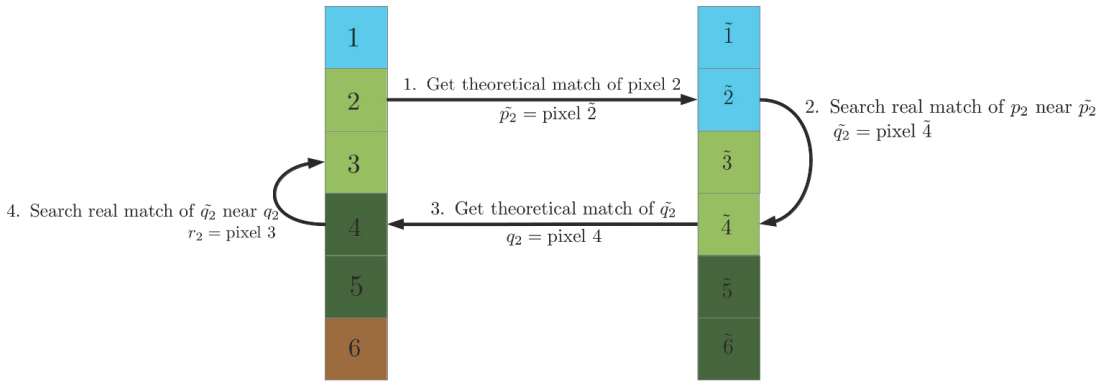

Figure 4.5: Schematic illustration of the correspondence search in the event of a un-successful retrieval of the corresponding pixels: r_2 is different from p_2, the pair $(p_2, \tilde{q}_2)$ is discarded.

pairs of matches. This is barely an issue for several reasons.

To start with, several quality estimation techniques exist over homogeneous areas [Loeb, 1997; Rao and Chen, 1999; Green et al., 2003; Six et al., 2004; Reulke and Weichelt, 2012]. Secondly, heterogeneous areas are the ones that are the most likely to be affected by sampling errors and adjacency effects [Mekler and Kaufman, 1980; Richter et al., 2006; Schläpfer et al., 2007]. Investigating a set of pairs where such elements are over-represented therefore implies retrieving a higher bound for uncertainties, hence producing a lower bound for imagery quality.

4.2.3 Note on Complexity

One could notice that the method remains comparably time consuming. For a window width of size w and a neighbourhood search distance of size n, each 'best match search' requires n^2w^2 spectral metric computations. The search complexity is therefore in $\mathcal{O}\left(Nw^2n^2\right)$ where N is the number of pixels of the scene.

For a typical set-up, one gets $w = n = 7$ and $N \approx 10^6$ pixels hence resulting in about 2.4 10^9 spectral distance evaluations. In the optimistic event, where 10^6 evaluation could be achieved by the computer each second, the overall computation time is of a magnitude of about 40 minutes. This theoretical result corresponds to computation times observed in practice.

In order to partially limit these delays, parallelism was introduced. Each pair of 'best matches' search is indeed independent from the others. The searches have therefore been paralleled using the OpenMP framework [OMP, 2012], hence resulting in a final complexity of $\mathcal{O}\left(Nw^2n^2/c\right)$ where c is the number of threads available for computations.

4.3 Parameters Choice

As presented in the previous section, the algorithm is based on three parameters:

- The neighbourhood size, n.
- The search window width, w.
- The spectral metric used for computations, $\mathcal{S}$

Although the window width is directly linked to the expected geo-location discrepancies between both flight-lines, the choice of the spectral metric and the optimal neighbourhood size are to be investigated in more details. This is done in section 4.3.1 and section 4.3.2 respectively.

4.3.1 Spectral Metrics

This sections aims at determining which of the six spectral metrics presented in section 2.4.1 are the most suitable for performing computations.

In order to find out which one is optimal for practical test-cases, the algorithm was run for the six metrics presented in section 2.4.1. All pixels from the first flight-line for which all six metrics found a match were studied. Since all metrics are investigating different aspects of the spectral discrepancies, it is unlikely that they are biased in the same way. In other words, if for a given pixel from the first flight-line, each metric is retrieving the same pair, then it is highly plausible that the pair is indeed the 'real' match. On the other hand, if all metrics but one are agreeing on a pair, it is likely that the latter is actually wrong.

This was in practice tested by running all six metrics on several data-sets before performing a majority voting on all pairs in which the pixel from the first flight-line got a match according to all six metrics. The pair getting the most votes is then considered as being the 'right' match and the corresponding metrics credited accordingly. Remaining matches are considered as erroneous. Results are presented in tables 4.1.

Metric	**Total**	**Validated**	**Errors**	**Time**
HySpex Kaufbeuren 2012: flight-lines $4A$ & $4B$: 171897 votes				
SA	1032602	142861	29036	2d 14 : 09
SCA	937288	144805	27092	5d 00 : 35
SGA	349734	120434	51463	5d 08 : 31
SID	N/A	N/A	N/A	$>$ 10days
l_2	1104184	129915	41982	1d 20 : 24
H_1	1115993	130054	41843	2d 23 : 27
HyMap Sokolov 2010: flight-lines 2 & 3: 19997 votes				
SA	90648	18473	1524	00 : 59
SCA	92502	18730	1267	04 : 06
SGA	74703	16343	3654	06 : 35
SID	89390	17410	1344	$\approx$ 2days
l_2	37435	12888	7109	00 : 34
H_1	38590	12936	7161	01 : 14

Table 4.1: Efficiencies of Metrics. $w = 11$, $n = 11$.

It should be noted that for 'big' data-sets, SID took too long for its execution. After ten days, it was stopped and discarded as being non usable for practical applications and the vote only took place among the remaining five metrics.

Besides SID, which is extremely slow compared to all other five metrics, only subtle differences are to be observed. To start with, in all cases, the spectral angle present a comparably good optimum between time of execution, total number of retrieved matches and amount of false matches. The classical metrics, namely H_1 and l_2 are very fast, especially the later. They perform quite well when the data is acquired using the same geometry (e.g. Kaufbeuren, flight-lines 4A-4B) but are comparably poor when the viewing angle differ (e.g. Sokolov). SCA and SGA are not offering tremendous advantages while taking slightly longer to run. SA is comparably good in all cases and is comparably fast. Since it is furthermore a widely used - if not the reference - metric within the hyperspectral community [Chang, 2000; Dennison et al., 2004; Garcia-Allende et al., 2008; Keshava and Mustard, 2002], it is therefore used in all the following computations.

These results are in compliance with finer manual comparisons of the SA and l_2 metrics made over the Sokolov 2010 scene [Müller, 2012].

From a more theoretical point of view, the apparent antagonism between SA and the l_2 metric should be noticed. The spectral angle performs better when pixels are imaged from different viewpoints. When both flight-lines are acquired with the same viewing geometry, the l_2 metric performs better. This can be explained as follow.

When both flight-lines are obtained from the same viewpoint, almost without contribution from the BRDF effects. As a result the variability consists mostly of an additive noise, which is not especially affecting the l_2 metric but has a comparable large influence on the spectral angle. On the other hand, when both view angles are differing, the BRDF effect is not negligible. Over a given neighbourhood, it can be seen as 'just' a multiplicative constant on the data. The spectral angle is therefore able do deal with this comparably easily but the l_2 metric is slightly less successful.

4.3.2 Neighbourhood Size

Now that a suitable metric - the spectral angle - has been identified, the optimum neighbourhood size has to be investigated. This is a slightly more complicated. As already illustrated on figure 4.3, a too wide neighbourhood might possibly endanger the retrieval step. On the other hand, a too small one is likely to lead to unreliable matches.

Assume one is using two consecutive and 'too small' neighbourhood distance. If both are failing to retrieve the right match for a given pixel, it is also likely that they fail the same way, since both are acting on almost the same area. The same remark also applies when working on 'too big' neighbourhood distances. It is therefore not possible to obtain an estimation of the corresponding optimal distance by a majority voting approach.

As a result, the focus was set on obtaining the minimum neighbourhood distance needed to obtain reliable results. Greater distances, even though potentially more stable, are unlikely to be used due to both the presence of potential distortions and the program execution complexity increase in $\mathcal{O}\left(n^2\right)$.

To achieve this goal a synthetic data was generated. A subset of the Kaufbeuren scene (see figure F.2) was perturbed by two sets of independent noise. The CPAM algorithm was then run between both synthetic scenes: see figure 4.6. Since this test-scene is actually presenting a comparably exhaustive list of land-covers (bare-soils, fields, forests, urban structures, ...) it is assumed that the results are representative of the general case.

Since both perturbed scenes come from the same data-set, the pairs of corresponding pixels are known beforehand. The algorithm results can therefore be directly checked: in other words

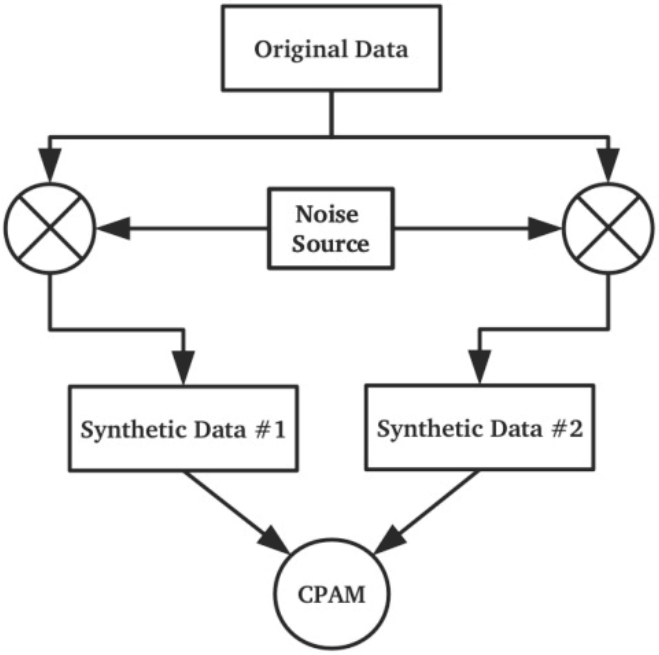

Figure 4.6: Schematic illustration of the synthetic scene generation steps.

$\tilde{q}_i$ should actually be $\tilde{p}_i$. The pixels not satisfying this condition are necessarily false matches.

The noise was chosen to be Gaussian with a standard deviation of 9% reflectance. It should be noted that this noise perturbation is applied on 'real' reflectance. For the third band in the visible domain of this data-set, reflectance values lie between 0% and 15%. A noise with a standard deviation of 9% reflectance is then equivalent to a noise with a *relative* standard deviation of 60%. The actual efficiency of the pair retrieval is therefore underestimated.

The search window size was set to $w = 7$. Since it is largely overshooting the actual discrepancies - $w = 0$ - this leaves room for a comparably wide range of false matches. The corresponding results are presented on figure 4.7. Since no geometric distortions were brought into the imagery, the amount of 'right matches' should be an ever increasing function of the neighbourhood distance. The goal is therefore 'only' to determine the minimum acceptable n, granting an acceptable ratio of right-matches.

Figure 4.7 reveals that neighbourhood sizes between 7 and 15 are granting a bad/good match ratio lower than 10%. Furthermore, for neighbourhood sizes greater than 10, more than the half of the input scenes are getting a valid match. For neighbourhood sizes greater than 15 pixels, only limited gains are retrieved. This tends to indicate that a suitable range for the neighbourhood size is $[\![7; 15]\!]$.

The next step would be to study the effect of such neighbourhood distances in terms of geometric distortions. To start with, the effect of a potential residual rotation is investigated. If an uncorrected rotation, θ, is present, it would yield over a displacement of d pixels an error of $d \tan \theta$ pixels in the orthogonal direction.

As indicated by table 4.2, choosing n ranging from 7 to 15 pixels are only marginally impacted by residual rotations - even in the very extreme case where these would reach 3 degrees [Schläpfer, 2010]. The produced offset is indeed always smaller than a pixel, hence limiting its impact on the matching process.

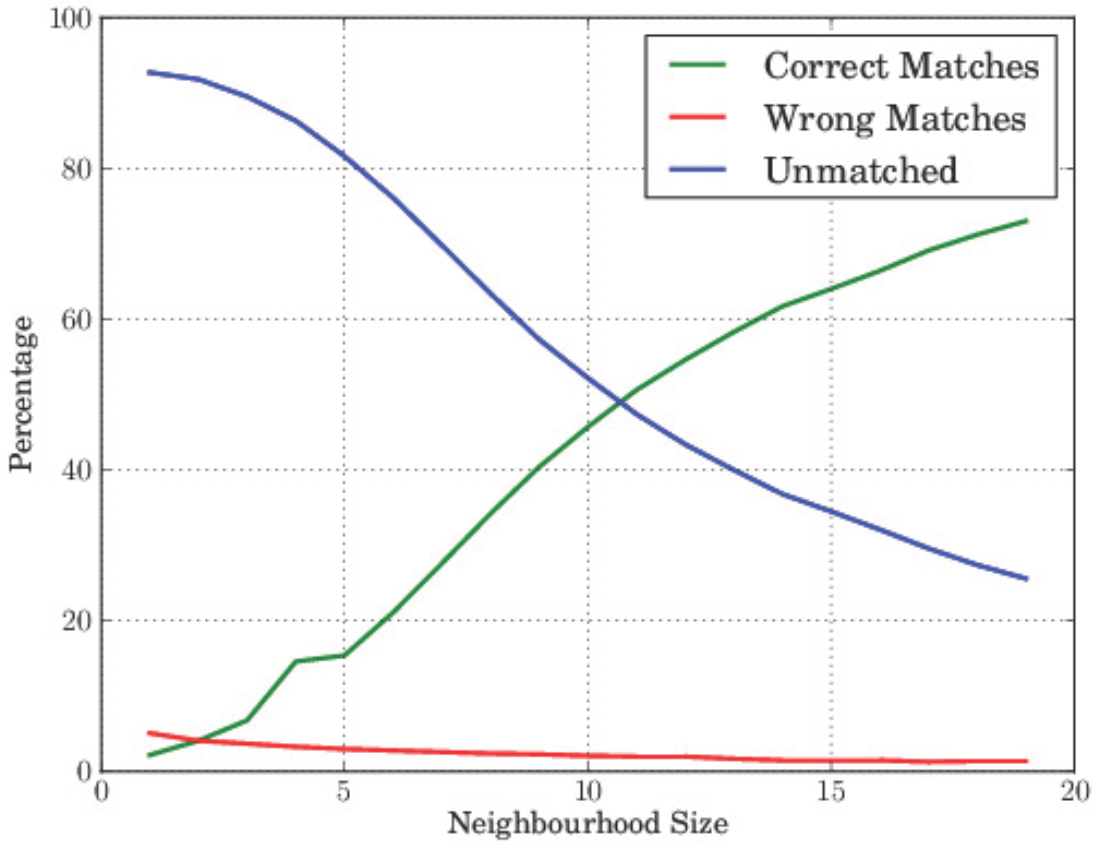

Figure 4.7: Effects of the neighbourhood size. Input scene: F.2, $w = 7$, with a Gaussian noise of standard-deviation of 9% reflectance.

		Distances (Pixels)		
		$d = 7$	$d = 10$	$d = 15$
Angles	$\theta = 1°$	0.12	0.17	0.26
	$\theta = 2°$	0.24	0.35	0.52
	$\theta = 3°$	0.37	0.52	0.79

Table 4.2: Effects of rotations on neighbouring sizes: residual rotations in terms of pixel's offsets (δR) as a function of distance - results expressed in pixels units.

The next step consists in checking the influence of an error associated to the digital elevation model. Although a global DEM bias does not affect the pixel pairs retrieval, relative errors within the DEM might in contrary disturb it. In the following, a local DEM relative error, ΔH, is investigated (see figure 4.8). The term relative error should be understood as being the discrepancy between the errors associated with two points from the DEM.

Using the Thales equality, one gets $\Delta H = \Delta D \cdot H/D$. This effect is at its worst at the flight-lines fringes, D is therefore set to be the half swath of the flight-line. H is set to the

flight altitude. This yields $D/H \approx tan(FOV/2)$. ΔD is of course corresponding to the size of one single pixel: this provides an estimate of the greatest ΔH which is leaving the matching process unaffected and FOV is the Field Of View of the sensor. It should finally be noted that, under nominal operations, the FOV is linked to the Ground Sampling Distance (GSD), this is the distance between two consecutive pixel footprints. In the following, GSD is considered to be an approximation of the pixel size.

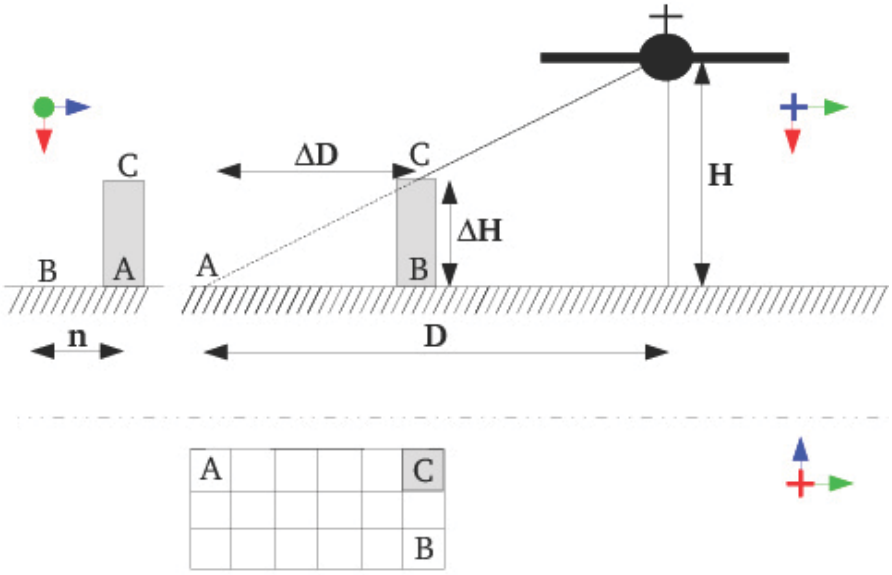

Figure 4.8: If the height ΔH is not accounted in the DEM, the spectral content of C will be placed in A. For CPAM results for pixel B to remain unaffected, ΔD has to be smaller than one pixel.

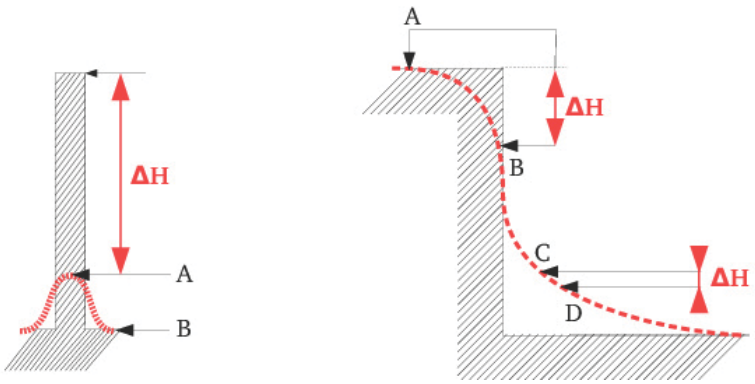

Figure 4.9: An illustration of the relative error influences for a smooth DEM (Red) on the terrain (black).

		FOV	GSD	D/H	ΔD (pixels)	ΔH (pixels)	G_7	G_{10}	G_{15}
Sensor	AHS (port 4)	90°	$5m$	1.0	1.0	1.0	14%	10%	6.7%
	AISA Dual	24°	$1m$	0.21	1.0	4.7	67%	47%	31%
	HyMap	61.3°	$4m$	0.59	1.0	2.0	29%	20%	16%
	HySpex (VNIR)	17°	$0.75m$	0.15	1.0	6.7	96%	67%	45%
	HySpex (SWIR)	14°	$1.5m$	0.12	1.0	8.1	116%	81%	50%
	HySpex (VNIR) x2	34°	$1.5m$	0.31	1.0	3.3	47%	33%	22%
	HySpex (SWIR) x2	28°	$3m$	0.25	1.0	4.0	57%	40%	26%

Table 4.3: Acceptable local DEM errors for CPAM. Ground sampling distance is purely indicative. HySpex is also investigated with its field of view expander (HySpex x2).

Table 4.3 presents some estimation of the maximum ΔH acceptable, depending on sensors typical set-up; grade slopes - as a function of a neighbourhood size n - are written G_n. The absolute admissible error is however of little use by itself if it is not linked with the distance over which the perturbations are occurring: see figure 4.9.

The DEM generation steps are generally including a 'smoothing' step [Arefi et al., 2011]: brutal local relative errors are therefore only occurring when the surface elevation itself is subject to brutal, well localised, changes (e.g. tower, ...) and are therefore restricted to a tiny subset of the complete scene. Even then, they have, for the investigated sensors, to be over about 4 metres high to have an impact: they are therefore ignored.

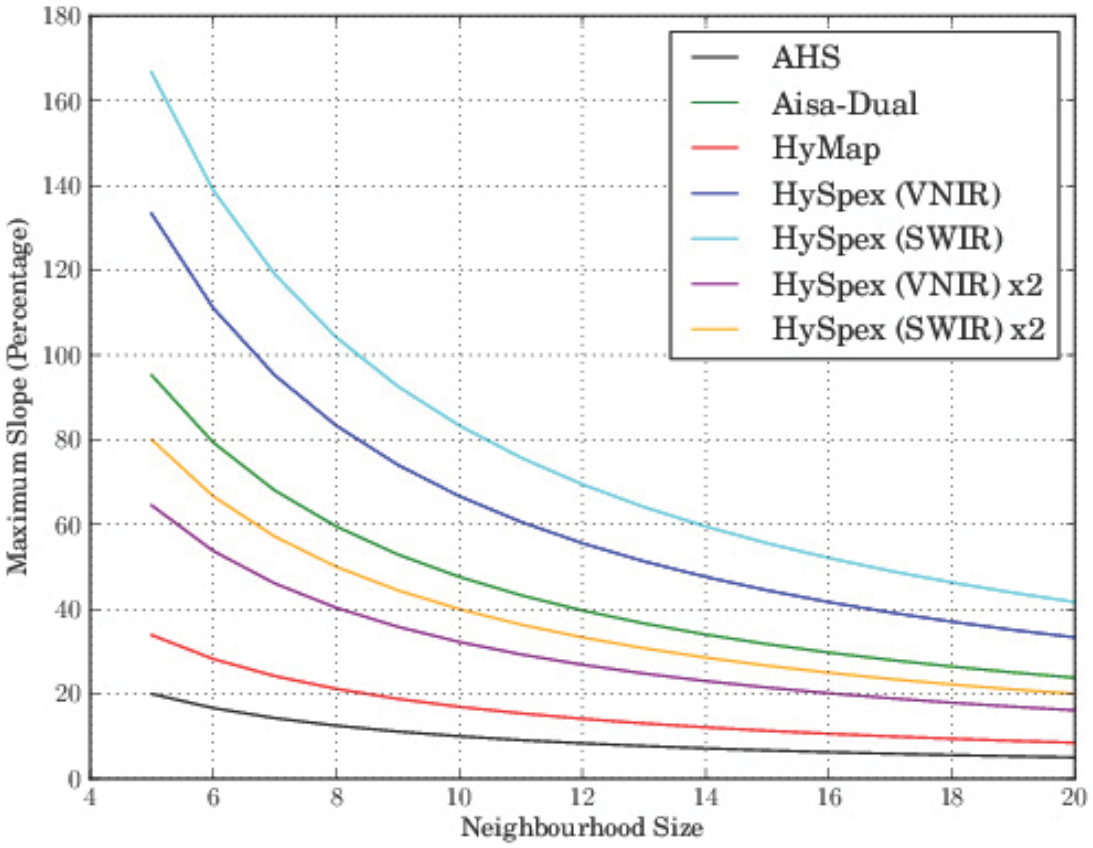

Figure 4.10: Maximum grade slopes of the elevation model errors as a function of the neighbourhood size.

On the other hand, gradual relative errors, as displayed on the right of figure 4.9 might have a strong effect. ΔH is therefore divided by the neighbourhood size, n, hence providing an estimation the corresponding maximal grade of slope of the DEM errors: $G_n =$. Detailed results are provided on figure 4.10.

According to Mukherjee et al. [2013], the relative slope errors when comparing a DEM extracted from Cartosat stereo-imagery with Shuttle Radar Topography Mission (SRTM, [Farr et al., 2007]) or ASTER based DEM are ranging from 12 to 15%. Dividing the SRTM absolute altitude error by the SRTM ground resolution at a latitude of 45° yields a similar upper bound: a grade

slope from about 14% [Farr et al., 2007].

These three estimations are however overshooting the actual relative slope error. The computations are indeed including biases present in the imagery, although such biases should discard themselves out when actually accessing the relative slope error.

Except for AHS, all investigated sensors are nevertheless tolerating relative DEM errors of up to 15% grade slope for neighbourhood sizes within 7 and 15 pixels. For AHS, values close to 7 should however be preferred, since these are the only ones accepting relative grade slopes error of up to 14%.

It should be noted that other geometric distortions might occur in the imagery (e.g. warping). In practice, they are far from being the predominant ones and do seldom occur: they have therefore not been investigated.

4.3.3 Conclusions

In order to retrieve optimum results, CPAM should be run with the spectral angle as a spectral metric and a neighbourhood size comprised between 7 and 15 pixels.

For sensors with very large field of views, neighbourhood sizes close to 7 should be preferred. On the other hand, for sensors with very small field of views, neighbourhood sizes greater than 15 could be used, as long as the maximum acceptable relative slope error stays over 15% (refer to figure 4.10). This does however imply a comparably large increase of the algorithm computational complexity.

When comparing scenes sharing the same viewing geometry, the l_2 metric can be considered as a slightly faster alternative to the spectral angle. It does however provide poorer results for 'classical' overlaps where the viewpoint differ. Other metrics are not bringing any noticeable advantages.

Finally the search window size has to be set-up manually, depending on the expected discrepancies in the scene. Choosing a too wide search distance is unlikely to hinder the matches retrieval, except if the scene is presenting periodic features. This, on the other hand, increases the computational time required to perform computations.

4.4 CPAM Validation

To start with, it should be noted that a first consistency-check step is performed by the algorithm itself. Indeed, the two steps approach only keeps pairs of pixels that are symmetrically corresponding, keeping only 'self-compliant' pairs, for which no error occurred or where errors cancelled themselves. Since the latter option seems unlikely in the event of non systematic uncertainties, they should correspond to a minority of events.

4.4.1 Illustrative Results

A second inspection can be done by analysing the amplitudes and directions of discrepancies between theoretical matches based on ortho-rectification and matches retrieved through CPAM: for example on figure 4.11. Although not properly being a validation step, this provides several points of information about the results.

First of all, there are comparably fewer matches retrieved over homogeneous areas. Such parts of imagery indeed present less spectral heterogeneity hence being less likely to retrieve two corresponding symmetric matches when running CPAM. This is worsened by the presence of the imagery variability (e.g. noise) which tends to further deprecate the spectral contrast between pixels, hence worsening the chances to get the right matches.

Secondly and in this particular example, a gradient of increasing discrepancies seems to occur from the top to the bottom of the subset. A finer analysis of the discrepancy directions reveals that they are perpendicular to the flight-line direction, hence indicating that these are likely due to a DEM error (see figure 2.4). Furthermore, when comparing the data - a mining front - with the digital elevation model used for ortho-rectification, it indicates that the discrepancies are indeed proportional to the expected DEM error. The latter is slightly outdated: its indicated 'bottom' is slightly off the actual one which coincides with the coal seams (see figure 4.12).

A few outliers, red pixels on figure 4.11, should at last be noticed. It is comparably hard to perform any assessment of their validity. They might indeed just be some errors made by CPAM or might correspond to an actual effect of the imagery such as:

- a moving object, or a vehicle, that got displaced between the two flight-lines acquisitions.
- a feature with a strong local altitude offset, which is not referenced on the DEM used for ortho-rectification: e.g. a tower. In the latter case the discrepancy should however be orthogonal to the flight direction.

This manual validation can of course be performed in a slightly more automated fashion, for example by computing the discrepancies histograms. If most of the discrepancies are located orthogonally to the flight-line directions, they are likely to be only due to DEM errors. Sections 4.4.2 and 4.4.3 present more qualitative assessment methodologies.

4.4.2 Validation on Synthetic Data

A major difficulty for obtaining an in-depth quantitative validation is however related to the presence of outliers (red pixels on figure 4.11) which might be due to both physical reasons and failures of the algorithm. Other issues are related to the heterogeneity of the discrepancies. As seen on figures 4.11 and 4.12, it might occur that discrepancies are varying over the imagery, hence making it a priori impossible to know which are the 'right' matches and which are not.

An approach similar to the estimation of optimum neighbourhood sizes was proposed: degrading an input image (figure F.2) by two sets of Gaussian noise of varying variances and searching the matches between both degraded images. The search was performed using the spectral angle

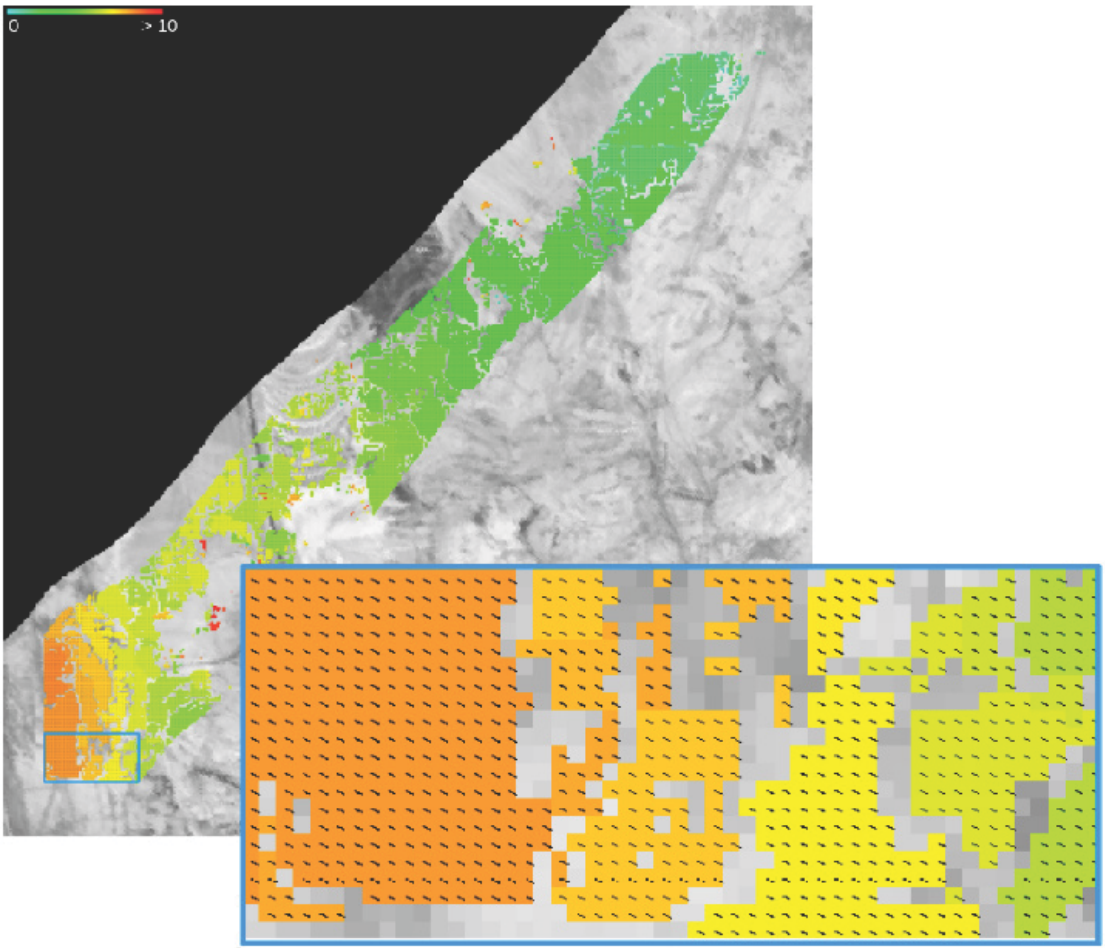

Figure 4.11: Retrieved pixels using CPAM on Sokolov-2010. Coloured pixels represent the found matches, their hue changing with the discrepancies between theory and retrieved matches. Black arrows stands for the discrepancies directions. Grey pixels represent original no-match pixels.

along with a neighbourhood $n = 7$ and a search window $w = 3$, hence granting a lower bound for the validity. Since both images are sharing the same geometry, retrieved matches for both of them should actually be the same pixel. If not, this indicates that an error occurred.

Results are presented on table 4.4. One can see that the algorithm stays stable until up 10% of signal-to-noise ratio, with a ratio of true matches by false matches below 10% of the retrieved matches. Note that, just as explained in section 4.3, this corresponds to a lower bound of the retrieved quality. These results tend to indicate that CPAM performs well over such land-covers, this is: a mixture of urban areas, open-fields, water bodies and forested areas.

Finally, although this computation is straightforward to perform on any single data-set, it only accounts for the modelled noise - in this case a Normal one. In order to get deeper understanding another complementary approach is presented in section 4.4.3.

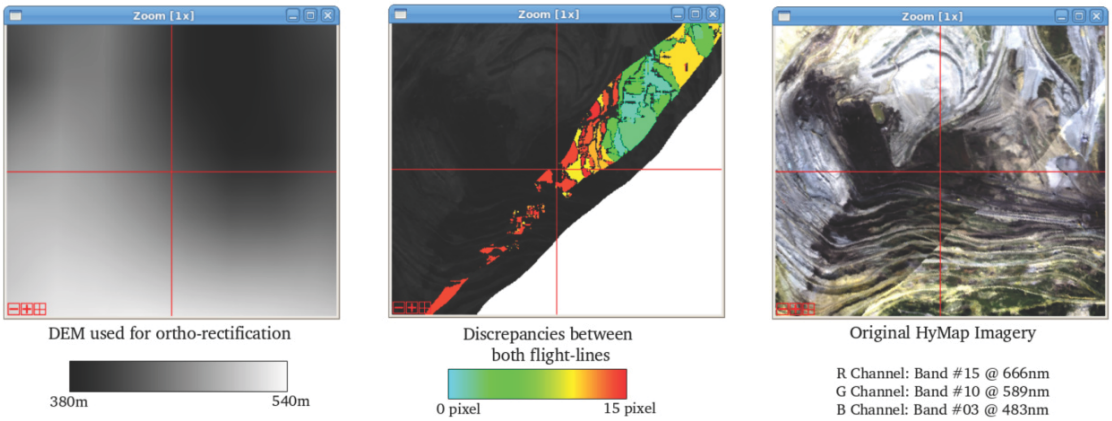

Figure 4.12: The coal seam - in black on the HyMap imagery, should correspond to the lowest part of the DEM. The latter is however slightly outdated, showing higher altitudes for the mining front and therefore leading to ortho-rectifications differences between both flight-lines.

Noise-Std	Correct	Wrong	Unmatched
0%	100%	0%	0%
1%	94.9%	0%	5.1%
2%	82.5%	0%	17.5%
3%	57.8%	0.3%	41.9%
4%	41.9%	0.9%	57.2%
5%	34.6%	1.4%	64.0%
6%	30.4%	1.7%	67.9%
7%	27.4%	1.9%	70.7%
8%	25.2%	2.1%	72.7%
9%	23.4%	2.3%	74.4%
10%	21.8%	2.5%	75.7%
15%	16.7%	3.5%	79.9%
20%	13.4%	5.1%	81.5%
25%	10.9%	6.8%	82.3%
30%	9.0%	8.7%	82.3%
40%	6.3%	13.1%	80.6%

Table 4.4: Percentages of correct and wrong matches found over the scene F.2 as a function of the noise standard-deviation (expressed in reflectance): $w = 3$, $n = 7$.

4.4.3 Triple Pass Validation

This validation approach is more comprehensive than the one based on synthetic data. On the other hand, it requires a specific data-set which might not be available, hence making it less handy for practical applications.

Principles

A more detailed validation step was performed using a triple pass validation. Given three flight-lines, A, B and C, CPAM can be run on each pair: $A-B$, $A-C$ and $B-C$. In the following, it is assumed that A B and C are sharing the same properties, including acquisition setup. If all retrieved matches are 'good' matches, then they should be compliant together. Each pixel from flight-line A (call them P_A) for which a match was retrieved in both B (P_B) and C (P_C) is investigated. If a match of P_B was found in flight-line C (Q_C), it is compared with P_C. If both are indeed the same pixel, it can be assumed that all three runs of CPAM were indeed successful. On the other hand, if a discrepancy occurred, there was at least an issue in one match retrieval (see figure 4.13).

The deduction that no error occurred as long $Q_C = P_C$ is of course not completely true. The case were two or more errors are actually cancelling themselves leads to the same result. This case is however very unlikely and can therefore be neglected in a first order approximation.

If the three flight-lines share the same acquisition conditions, it can be assumed that the proportion of 'good' match retrievals, o_0, is the same for all three CPAM runs. The same is valid for the proportion of wrong matches, e_0 and the proportion of pixels that did not get any matches v_0.

These quantities can in turn be linked to the global proportions of good, wrong and absence of matches: respectively O, E and V - at least as long as the proportion of false matches - o_f - leading to $P_C = Q_C$ can be neglected:

$$\begin{cases} O &= o_0^3 + o_f \approx o_0^3 \\ E &= e_0^3 + 3o_0e_0^2 + 3o_0^2e_0 \\ V &= v_0^3 + 3(o_0+e_0)v_0^2 + 3(o_0+e_0)^2v_0 \end{cases} \tag{4.4.1}$$

Solving for the first line is straightforward. This leads to a third order equation in e_0 for the second line. Which can in turn be solved, yielding three potential solutions for e_0. The 'right' solution can then be identified using requirements for e_0 to be a proportion - namely that is has to be a purely real number comprised between zero and one.

The third line can finally be transformed by means of the necessary condition $O+V+E=1$, which leads back to the requirement $o_0+v_0+e_0=1$ through equation 4.4.2, hence $v = 1-o_0-e_0$, which could also have been inferred directly form the probabilistic interpretation of o_0, e_0 and v_0.

$$\begin{aligned} 1 &= O+V+E \\ &= o_0^3+e_0^3+v_0^3+3o_0e_0^2+3o_0^2e_0+3o_0v_0^2+3e_0v_0^2+3o_0^2v_0+3e_0^2v_0+6o_0e_0v_0 \\ &= (o_0+e_0+v_0)^3 \end{aligned} \tag{4.4.2}$$

$$\Rightarrow \quad o_0+e_0+v_0=1$$

When only two flight-lines are sharing the same acquisition set-up, the proportion of right matches is slightly more difficult to assess. To start with, it is assumed that B and C are the two flight-lines sharing the same set-up. This means that CPAM runs on $A-B$ and $A-C$ have the same proportion of 'right' matches o_1, while CPAM runs on $B-C$ are associated with a proportion of rightful matches of o_2. This leads to the ill posed problem $o_1^2o_2 = O$ which can only be solved if an estimation of o_2 - or o_1 - has already been done. Noting similarly e_1 and e_2 the error rates and v_1 or v_2 the non match rates, equation 4.4.1 becomes equation 4.4.3.

$$\begin{cases} O &= o_1^2 o_2 \\ E &= e_1^2 (e_2 + o_2) + 2e_1 o_1 (e_2 + o_2) + e_2 o_1^2 \\ V &= (e_2 + o_2) \left(2v_1 (e_1 + o_1) + v_1^2\right) + v_2 (o_1 + v_1 + e_1)^2 \\ &= (e_2 + o_2) \left(2v_1 (e_1 + o_1) + v_1^2\right) + v_2 \end{cases} \tag{4.4.3}$$

As long as the the triplet (o_2, e_2, v_2) is known - solving for (o_1, e_1, v_1) is straight forward using the above mentioned requirements. In practice, estimation of (o_2, e_2, v_2) is based on solutions from equation 4.4.1: namely (o_0, e_0, v_0).

As before, the first line directly provides o_1. The second equation is simply a second order equation in e_1, see equation 4.4.4. Since all terms are positive, $E \geqslant e_0 o_1^2$ which leads to a determinant - Δ - not only positive but also greater than the squared first order coefficient: $4 (o_1 (e_0 + o_0))^2$. Since e_1 has to be positive, the only acceptable solution is the one where the discriminant Δ is added:

$$e_1^2 (e_0 + o_0) + 2e_1 o_1 (e_0 + o_0) + e_0 o_1^2 - E = 0$$

$$\Delta = 4 \left[o_1^2 (e_0 + o_0)^2 + (e_0 + o_0) \left(E - e_0 o_1^2\right)\right] \geqslant 4 \left[o_1 (e_0 + o_0)\right]^2 \geqslant 0$$

$$\Rightarrow \begin{cases} e_1 &= \frac{0.5\sqrt{\Delta} - o_1 (e_0 + o_0)}{e_0 + o_0} \geqslant 0 \\ \text{or} & \\ e_1 &= -\left(\frac{o_1 (e_0 + o_0) + 0.5\sqrt{\Delta}}{e_0 + o_0}\right) \leqslant 0 \end{cases} \tag{4.4.4}$$

Finally, the third equation - providing v_1 - can advantageously be replaced by $v_1 = 1 - o_1 - e_1$, being straight forward to solve.

This theoretical framework requires two adaptations before being put in practice. If one flight-line, say C, did not cover a specific location: only matches between A and B can be retrieved on this specific spot. This does however not correspond to a lack of match between A and C or B and C. A first pass is therefore made to discard these points from comparison based on their distance to image boundaries and their theoretical geo-location.

A second issue concerns the slight shifts between the pixels' footprints of different flight-lines. In practice P_A, P_B, P_C and Q_C are not exactly overlapping. It might even be the case that P_C and Q_C are not exactly the same pixel, although all matches retrievals performed perfectly (e.g. see figure 4.13). When this is the case, P_C and Q_C do however have to be immediate neighbour due to the density of matching. The comparison therefore has to take place with a a tolerance of ± 1 pixel.

Application

The approach was applied on the Kaufbeuren 2012 data-set F.1 (CPAM run with $w = 10$, $n = 10$). This scene presents a comparably high variability of land cover (urban, forests, fields...) presenting both homogeneous and heterogeneous areas. It does moreover have three flight-lines sharing almost exactly the same acquisition set-up: $4A$, $4B$ and $4C$, hence allowing for a complete analysis.

The first case, where all flight-lines are sharing the same acquisition set-up corresponds to an investigation of flight-lines $4A$ and $4B$, $4A$ and $4C$, $4B$ and $4C$: 4761244 pairs were retrieved between flight-lines $4A$-$4B$. Amongst them, 3507385 were also corresponding to a match between $4B$ and $4C$:

- 88082 of them being associated with errors.

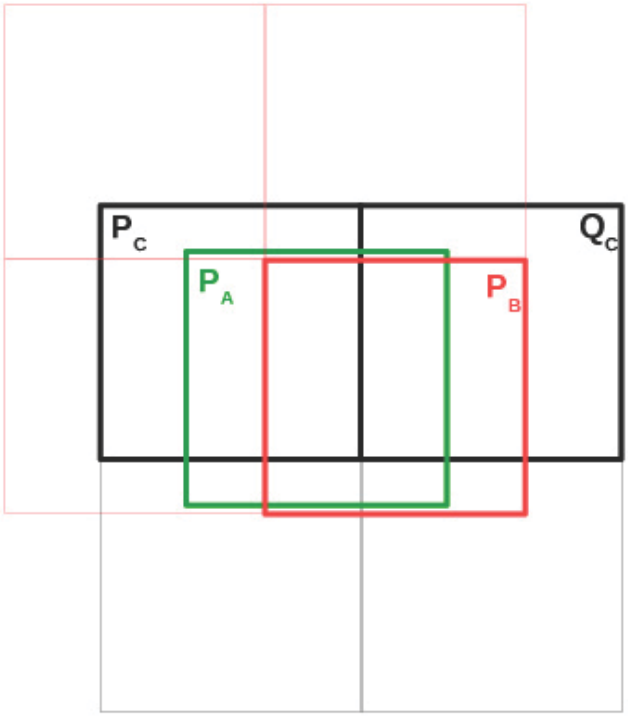

Figure 4.13: Illustration of a pathological triple match where due to different pixels footprints offsets P_C and Q_C differ. Green: footprints of pixels from flight-line A, red: pixels from flight-line B, black: pixels from flight-line C.

- 3419303 of them being compliant together: $P_C = Q_C$.

Which yields the following results for the VNIR imagery:

- Proportion of correct matches $o_0 = 89.55\%$
- Proportion of false matches $e_0 = 0.76\%$
- Proportion of absence of matches $v_0 = 9.69\%$

Results for the SWIR detector are slightly less good, especially in terms of numbers of unmatched pixels. This is due to the lower signal to noise ratio in the SWIR domain.

- Proportion of correct matches $o_0 = 74.03\%$
- Proportion of false matches $e_0 = 2.33\%$
- Proportion of absence of matches $v_0 = 23.64\%$

These results tend to indicate that the ratio of good/wrong matches is largely greater than 30 to 1, hence indicating the method validity. It should on the other hand not be forgotten, that the exact figures are scene dependent. For example when taking the extreme case study of a purely homogeneous surface, one likely gets very different results. On the other hand, since this

Kaufbeuren scene is presenting a comparably wide range of land-cover types, these results would be considered as being '*average*' values, average being here understood in its qualitative sense only.

Setting the previous results as being the first term, $(o_2 = o_0, e_2 = e_0, v_2 = v_0)$ and solving equation 4.4.3, it is then possible to gain estimates of the matching accuracy between flight-lines 3, 5, 7 and 8 and their $4A$, $4B$ and $4C$ counterparts in the Kaufbeuren 2012 data-set: these results are summarised in 4.5 and 4.6.

Triplet of flight-lines	**Triplet Errors**	**Triplet Valid Matches**	**Triplet Uncompleted**	**Total Pixel Number**
$4A-4B-4C$	$e_0 = 0.76\%$	$o_0 = 89.55\%$	$v_0 = 9.69\%$	4 761 244
$3-4A-4B$	$e_1 = 2.57\%$	$o_1 = 73.25\%$	$v_1 = 24.18\%$	927 303
$3-4A-4C$	$e_1 = 2.40\%$	$o_1 = 76.49\%$	$v_1 = 21.11\%$	1 216 531
$3-4B-4C$	$e_1 = 3.20\%$	$o_1 = 71.51\%$	$v_1 = 25.29\%$	971 228
$5-4A-4B$	$e_1 = 2.65\%$	$o_1 = 72.30\%$	$v_1 = 25.05\%$	749 939
$5-4A-4C$	$e_1 = 1.29\%$	$o_1 = 76.10\%$	$v_1 = 22.61\%$	645 664
$5-4B-4C$	$e_1 = 2.83\%$	$o_1 = 70.67\%$	$v_1 = 26.50\%$	773 077
$7-4A-4B$	$e_1 = 0.82\%$	$o_1 = 85.84\%$	$v_1 = 13.33\%$	1 593 269
$7-4A-4C$	$e_1 = 0.91\%$	$o_1 = 87.24\%$	$v_1 = 11.85\%$	1 642 237
$7-4B-4C$	$e_1 = 1.46\%$	$o_1 = 82.03\%$	$v_1 = 16.50\%$	1 624 031
$8-4A-4B$	$e_1 = 1.57\%$	$o_1 = 81.76\%$	$v_1 = 16.67\%$	905 003
$8-4A-4C$	$e_1 = 1.58\%$	$o_1 = 82.73\%$	$v_1 = 15.69\%$	955 576
$8-4B-4C$	$e_1 = 2.44\%$	$o_1 = 76.06\%$	$v_1 = 21.50\%$	951 154

Table 4.5: Results of the triplet analysis over Kaufbeuren, VNIR detector: $n = 10$, $w = 10$.

Triplet of flight-lines	**Triplet Errors**	**Triplet Valid Matches**	**Triplet Incompleted**	**Total Pixel Number**
$4A-4B-4C$	$e_0 = 2.33\%$	$o_0 = 74.03\%$	$v_0 = 23.64\%$	586 522
$3-4A-4B$	$e_1 = 0.19\%$	$o_1 = 48.76\%$	$v_1 = 51.05\%$	29 396
$3-4A-4C$	$e_1 = 1.18\%$	$o_1 = 61.72\%$	$v_1 = 37.09\%$	23 816
$3-4B-4C$	$e_1 = 0.12\%$	$o_1 = 49.23\%$	$v_1 = 50.65\%$	32 296
$5-4A-4B$	$e_1 = 2.29\%$	$o_1 = 20.56\%$	$v_1 = 77.15\%$	14 694
$5-4A-4C$	$e_1 = 3.20\%$	$o_1 = 52.92\%$	$v_1 = 43.88\%$	54 056
$5-4B-4C$	$e_1 = 2.90\%$	$o_1 = 20.07\%$	$v_1 = 77.03\%$	15 460
$7-4A-4B$	$e_1 = 2.85\%$	$o_1 = 66.67\%$	$v_1 = 30.48\%$	147 485
$7-4A-4C$	$e_1 = 2.09\%$	$o_1 = 73.66\%$	$v_1 = 24.25\%$	158 719
$7-4B-4C$	$e_1 = 3.08\%$	$o_1 = 66.74\%$	$v_1 = 30.17\%$	148 735
$8-4A-4B$	$e_1 = 5.46\%$	$o_1 = 55.82\%$	$v_1 = 38.72\%$	76 358
$8-4A-4C$	$e_1 = 4.63\%$	$o_1 = 60.83\%$	$v_1 = 35.54\%$	81 176
$8-4B-4C$	$e_1 = 5.50\%$	$o_1 = 55.80\%$	$v_1 = 38.70\%$	77 458

Table 4.6: Results of the triplet analysis over Kaufbeuren, SWIR detector: $n = 10$, $w = 10$.

To start with, one can notice that the results over the SWIR detector are slightly worse than the results obtained with the VNIR detector. Reason for this is likely the higher signal to noise ratio obtained toward the infra-red part of the spectrum.

The comparably low score obtained in the SWIR domain between flight-lines 5, $4A$ and $4B$ or 5, $4B$ and $4C$ can be explained by the comparably low number of pixels present in the overlapping parts - e.g. less than 150 false matches - which make any exploitation of the corresponding percentages uncertain.

In all cases the good-to-bad ratio is above 10 and in some occasions, well above 100 which indicates that the method is suitable for further analysis.

Conclusion

The triple pass validation has shown that - over the Kaufbeuren data-set - CPAM produced valid results. Since this area contains both urban settlements, forested areas, open-fields as well as some water bodies, it can be assumed that these results are representative for areas presenting similar land-covers.

Furthermore, and due to the comparable high number of large homogeneous fields in the southern part of the scene one can consider that the scene is comparably homogeneous, thus providing for a lower bound of the matching quality. This is further stressed by the use of only 'half' spectra. Since the data is provided as two subsets, VNIR and SWIR, computations are not made on the complete spectrum, hence discarding some of its information and leading to a underestimation of the actual potential.

Finally, it should be noticed that the comparison of any flight-line with respect to flight-lines $4A$ and $4C$ is actually better than comparisons between $4A$ and $4B$ or $4B$ and $4C$. A closer examination of the set-up indicates that the flight-line $4B$ is actually slightly offset to the east hence modifying slightly the results. This yields, writing as before o_0 the rate correct matches for two similar flight-lines and o_1 this rate for two dissimilar flight-lines:

$$o_{4A\to 4B} = o_{4B\to 4C} < o_0 < o_{4A\to 4C}$$

When solving for o_1, this yields in turn under-estimations (resp. over-estimations) for $o_1\,(4A \to 4C)$ (resp. $e_1\,(4A \to 4C)$ and $v_1\,(4A \to 4C)$. On the contrary

$$o_1\,(4B \to 4C) = o_1\,(4A \to 4B)$$

while the corresponding error-rates and un-match rate are overestimated. The results are therefore to be understood as providing boundaries for e_1, v_1 and o_1.

Results obtained when comparing flight-lines $4A-4B$ and $4B-4C$ to another one are however stable which indicates the method validity. In any case, the retrieved quality was usually better than a 10 to 1 ratio between good and erroneous matches. This confirms the method suitability.

At last, it should be stressed that although this method allows for an in-depth estimation of the algorithm accuracy, it requires a specific data-set with three overlapping flight-lines. Such a set-up is rarely available, hence weighting the method's practical utility.

4.5 CPAM Results and Analysis

This section presents several differing applications of the CPAM tool to different data-sets along with an analysis of the corresponding results. For all imagery snapshots, coloured pixels indicate retrieved matches, black and white ones representing the original data pixels for which no match could be retrieved. The hue is directly representing the geo-location discrepancies between both flight-lines. Cyan represent no discrepancies, red discrepancies of $\sqrt{2}w$ pixels.

4.5.1 HyMap

Sokolov 2010

The run of CPAM on each pair of flight-lines from the data-set acquired over Sokolov in 2010 led to the retrieval of a total of 304750 matches with a neighbourhood size of $n = 7$ and search window size of 10.

The distribution of the geometric discrepancies between each pairs of flight lines are given in figure 4.15. An interesting feature is the alignment structure they present. Under the approximation that columns are heading south and that rows are heading east, the headings of these structures were computed and compared to the flight-lines heading (see table 4.7). In most cases the retrieved heading is roughly orthogonal to the flight-heading, hence indicating that most discrepancies are indeed due to outdated DEM information. It should however be noted that this can only be seen as a 'qualitative' validation.

Given a flight set up made of two perfectly parallel flight-lines, a DEM error produces a geo-location discrepancy orthogonal to the current headings. On the other hand, the aircraft headings vary during the flight-line acquisitions and both flight-lines are never completely parallel. Aggregating all geometric discrepancies into a histogram therefore mixes them, hence not necessarily resulting in a pure orthogonal result.

Another point is the line fitting. When the DEM is outdated in a heterogeneous fashion, discrepancies are indeed spread over the line. This can nicely be illustrated between flight-lines 2 and 3 which encompass the active mining area. On the other hand, if the DEM error is merely constant over the scene, geometric discrepancies are spread over a single point, hence making it difficult to retrieve the 'direction'. This is for example the case between flight-lines 4 and 5 encompassing areas with little heterogeneity in the underlying DEM.

Another interesting feature is the accumulation points on the borders when plotting the discrepancies histogram between flight-lines 2 and 3. Such a behaviour is likely to be linked with a too small search window. When this is the case, the actual correspondence can not be retrieved. A 'close' correspondence however might be retrieved near the actual one: this is near the border. When relaunching the computation with a larger search window, $w = 20$, this effect is indeed discarded (see figure 4.14). In the following and unless specified otherwise, these flight-lines are processed with these values, when others are processed with $n = 7$ and $w = 10$.

Flight-lines	Flight-line Heading	Discrepancies Heading	Differences
1	38	121	83
2	38		
2	38	128	90
3	38		
3	38	119	81
4	219		
4	219	97	122
5	220		
5	220	119	101
6	220		
6	220	103	63
7	38		

Table 4.7: Comparison between the flight-line headings and the main-directions of the discrepancies: example of Sokolov 2010. Heading are expressed as angles (degrees) expressed clockwise from north.

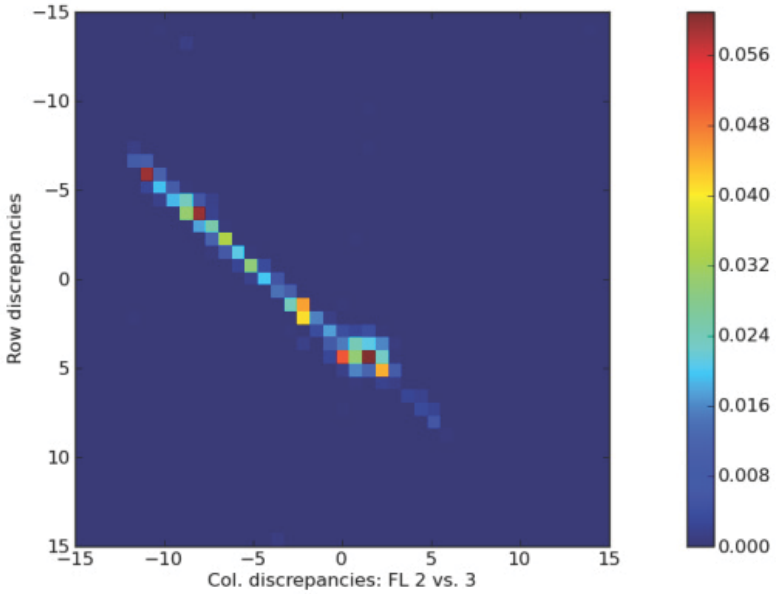

Figure 4.14: Histogram of geometric discrepancies for the Sokolov 2010 data-set (see F.2.1): new execution for flight-lines 2 and 3 with $n = 7$, $w = 20$.

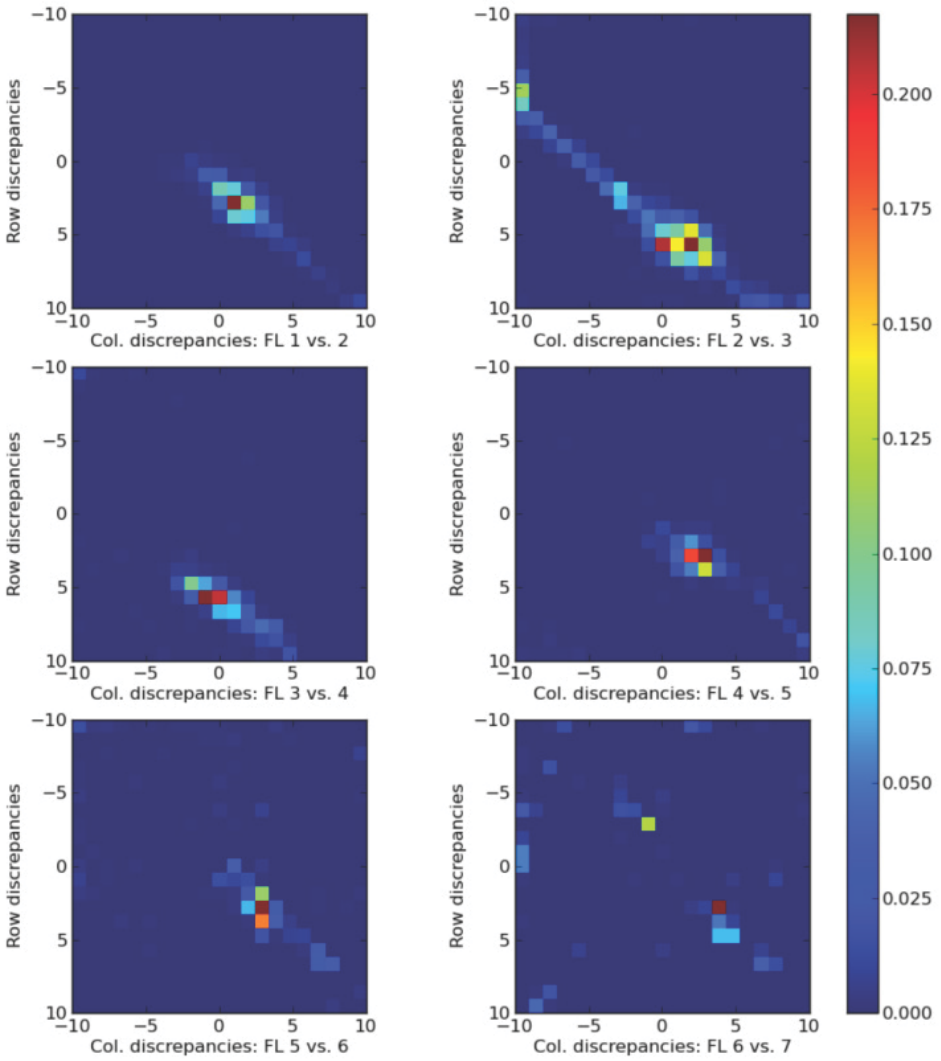

Figure 4.15: Histograms of geometric discrepancies for the Sokolov 2010 data-set (see F.2.1): $n = 7$, $w = 10$.

4.5.2 HySpex - DLR

Data produced by the dual HySpex instrument as operated by DLR G.4 was tested over two test sites. To start with, tests performed over the DLR calibration site - Kaufbeuren (see F.1) - are presented in section 4.5.2, and correspond to the ones used for the validation steps.

Kaufbeuren 2012

The results for the SWIR detector are presented in figure 4.16, where several properties observed in the previous sections can be retrieved.

To start with, one can notice the comparably noisy results obtained between flight-lines 5 and $4B$ which are due to a comparably small amount of matches that were retrieved. Comparisons between flight-lines $4A$, $4B$ and $4C$ present almost no dissimilarities: since they share similar viewpoints, DEM errors have little or no effect on them. This is especially true for the pair $4A - 4C$ which are sharing almost exactly the same acquisition geometry.

Finally comparisons between flight-lines $7 - 8$ and flight-lines $4A - 4B - 4C$ which are not parallel reveal interesting fine cross patterns. Since each branch is actually orthogonal to one flight-line, these are likely linked to slight issues within the underlying DEM. Their straightness however tends to indicate that such issues are not affecting the data very much.

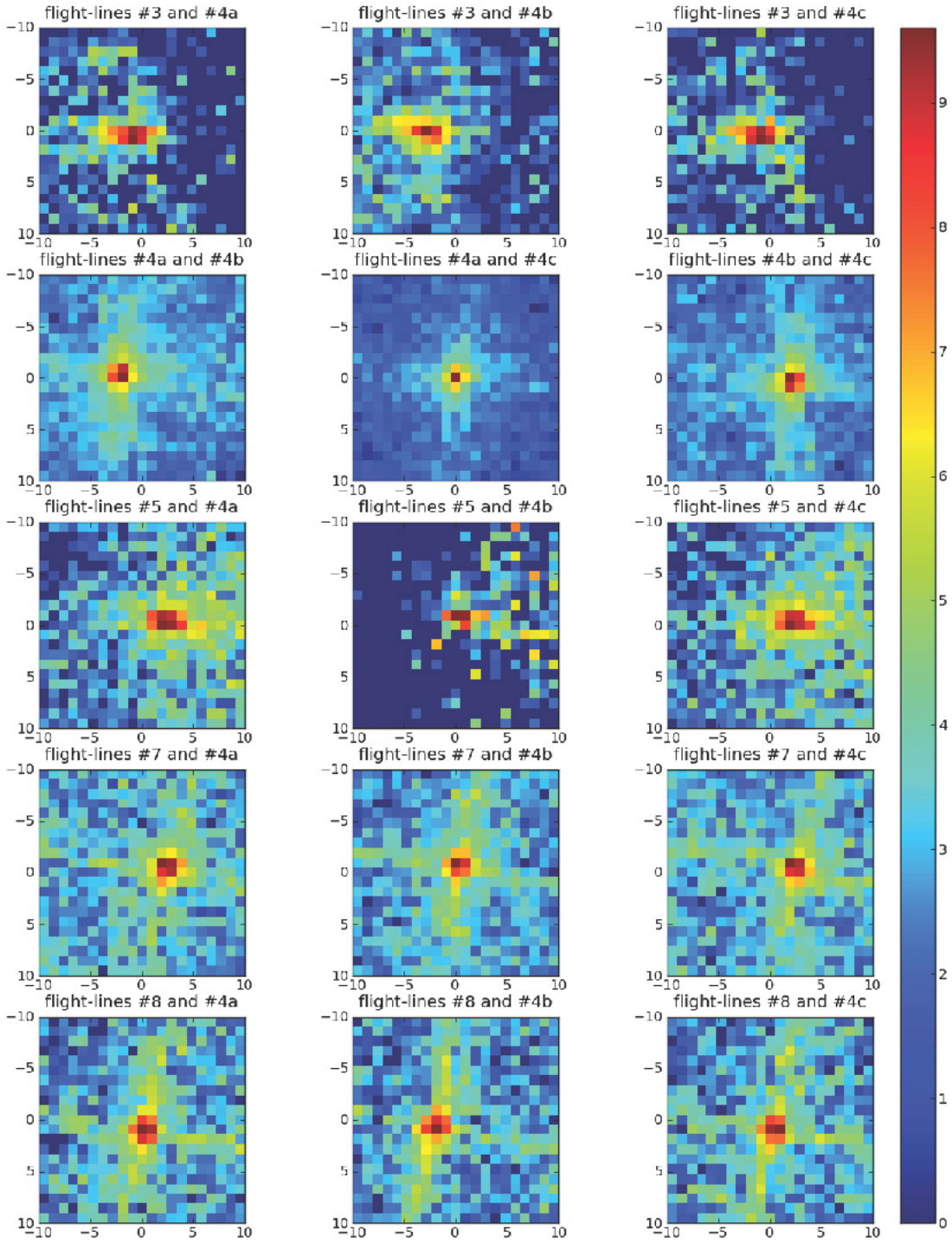

Figure 4.16: Logarithmic histograms of geometric discrepancies for the Kaufbeuren 2012 dataset: SWIR detector, $w = 10$, $n = 10$.

Sokolov 2013

The HySpex sensor was furthermore used to provide imagery over the Sokolov test site in 2013 (see section F.2.3). CPAM was also successfully run on the data-set providing a dense matching for both VNIR and SWIR parts of the detector, similar to the results obtained with HyMap.

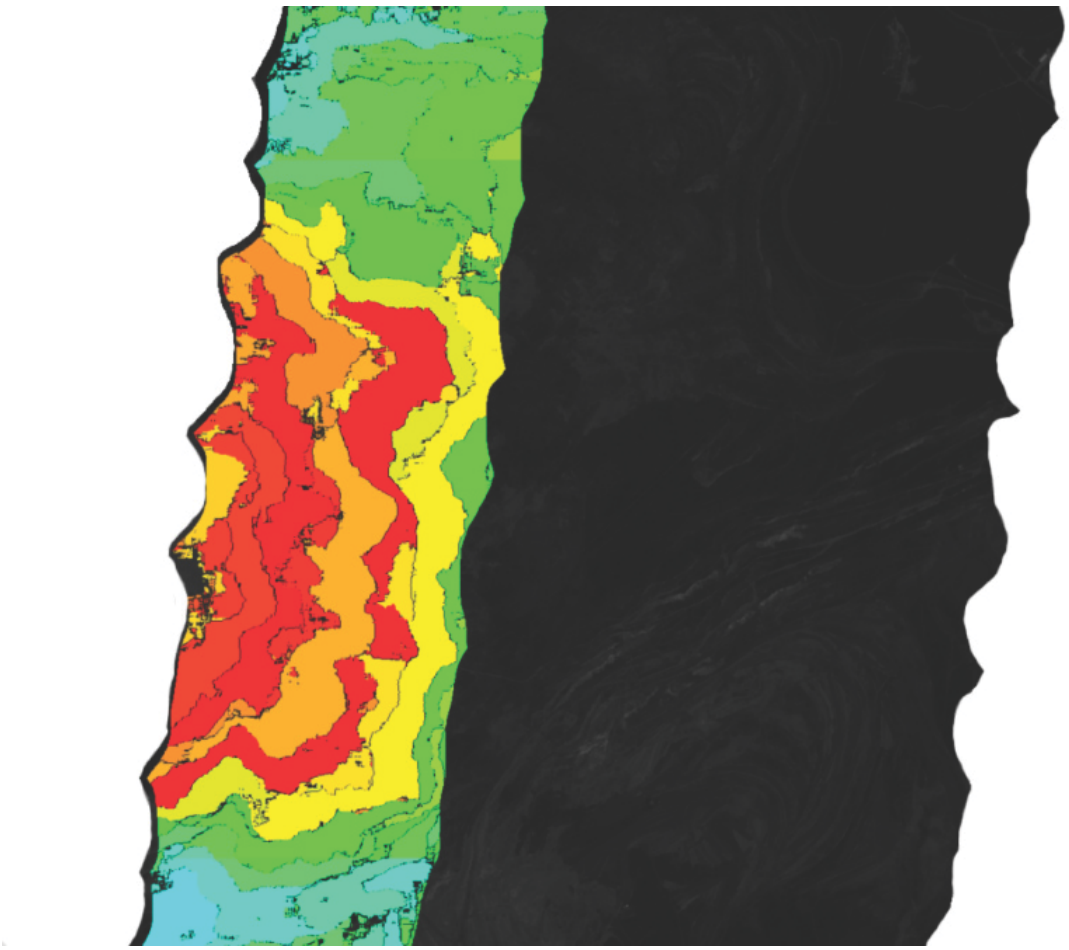

Figure 4.17: Illustrative result for Sokolov 2013 data-set: subset of matches retrieved between flight-lines 5 to 6 - VNIR $w = 10$, $n = 10$.

A particularity of this data-set is its varying pixel sizes. For the VNIR part of the detector, these are ranging from 2.38 meters to 2.44 meters[2]. Although the matching is comparably unaffected, this leads to an interesting chessboard pattern: see figure 4.18. On this specific example, pixels from flight-line 9 are 2.427 meters wide, while the ones from flight-line 17 are 2.390 meters wide. This yields that a distance of 65 pixels in flight-line 9 actually corresponds to 66 pixels in flight-line 17. This does in turn imply that over a period of 65 pixels, the matching has a discontinuity of 1 pixel, hence leading to the observed pattern.

[2]The reason for this lies in an optimal size computation, obtained from the relative altitude with respect to ground. Due to terrain variability, the same absolute flight altitude leads to different relative altitude, thus differing pixel's sizes.

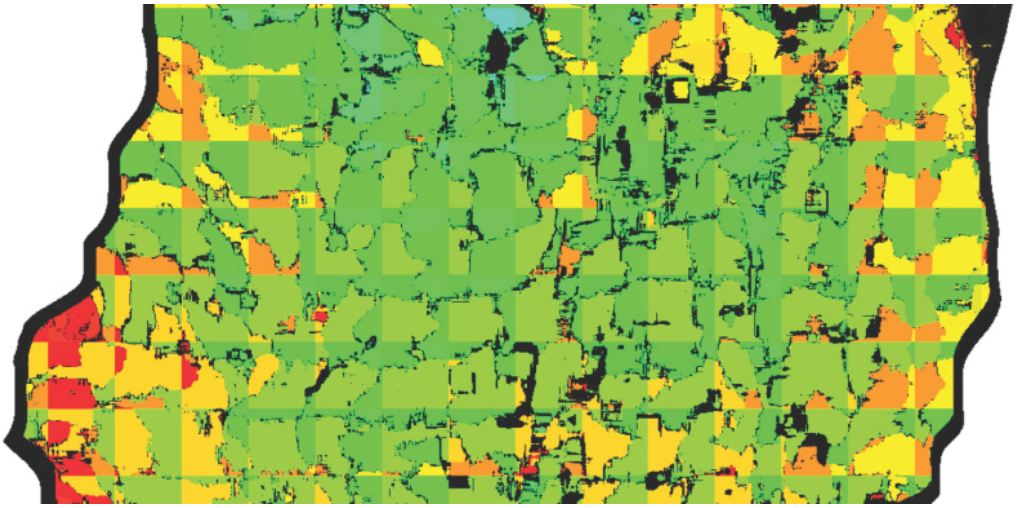

Figure 4.18: Tiling issues on the Sokolov 2013 data-set: flight-lines 9 to 17 - VNIR $w = 10$, $n = 10$.

4.5.3 AISA

Finally the algorithm was run on another hyper-spectral airborne data-set. An AISA data-set acquired over the Witbank municipality in South Africa. The scene differs in several points to the Kaufbeuren and Sokolov sites in the following sense:

- It is a semi-arid landscape, slightly differing from the Central-European data-sets.
- It was flown with a high accuracy LiDAR survey which provided an up to date digital elevation model for ortho-rectification.
- There are many subsidence structures (small to large scale) along with natural relief.

This was therefore used to check the algorithm validity under different conditions. Since the DEM and therefore the ortho-rectification is up-to-date, only small discrepancies are to be expected, the search window size was therefore set to a comparably small size of $w = 3$. The neighbourhood was set to 10. The corresponding histograms for the first 60 flight-lines.One can note that the match density is also very high (e.g. between flight-lines 14 and 15: see figure 4.20). Several remarks should be made.

To start with, the discrepancies are indeed largely spread along an horizontal axis, hence corresponding to an east-west direction - orthogonal to flight headings. This corresponds to slight DEM related errors, typically due to the discrepancies between the surface model (including roofs, canopy, etc) and the elevation model used for rectification which is only accounting for the terrain.

It should however be noted that in some cases the vertical (i.e. along flight-directions) discrepancies are greater than the search window. This is typically the case between flight-lines 30 and 31: fifth row and last column on figure 4.19. This corresponds to vertical discrepancies between both images, likely due to a timer drift in the acquisition system (see figure 4.21).

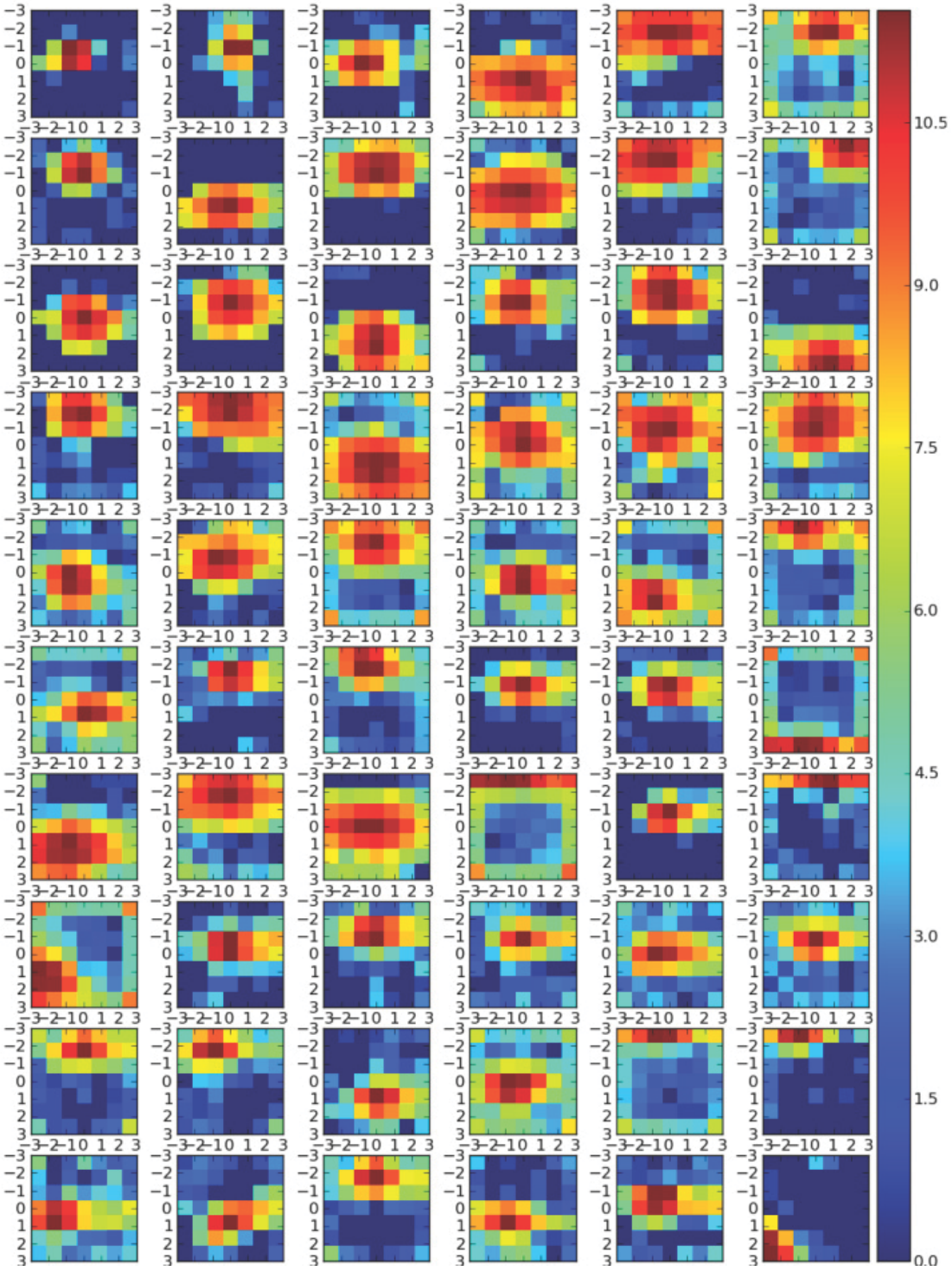

Figure 4.19: Logarithmic histograms of geometric discrepancies for the Witbank 2013 data-set (see F.3): First 60 flight-lines, $w = 3$, $n = 10$.

Figure 4.20: Illustration of the matching density over the Witbank 2013 data-set: example of flight-lines 14 and 15 with $w = 3$, $n = 10$. Coloured pixels are the retrieved matches, their hue changing with discrepancy intensity.

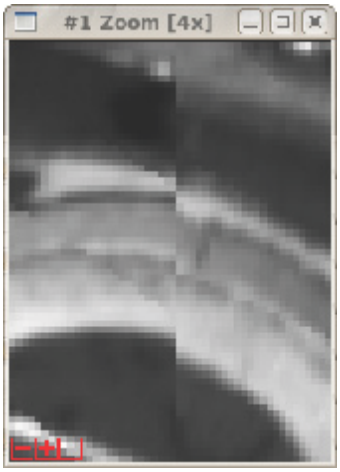

Figure 4.21: Observed vertical discrepancies between flight-lines 30 and 31 of the Witbank 2013 data-set.

4.5.4 Extension to Satellite Data

Despite having been designed for airborne data-sets, CPAM can also be successfully applied to spaceborne multi- and hyperspectral sensors. The only requirement applies to the satellite imaging capacities: it has to allow the acquisition of two overlapping data-sets under almost the same condition, this is over an interval of time as little as possible.

The simplest way to achieve this is to take advantage of along-track de-pointing capacities of some satellites. Such satellites are able to acquire image 'in front of them' and/or 'behind them'. This makes it possible to obtain data shortly before over-passing the scene as well as shortly after. The time difference between both acquisitions is typically of a few minutes, granting similar acquisition conditions. Such an de-pointing capacity is relatively common among satellites with stereo-imaging capacities.

One of those is the World-View-II satellite operated by Digital-Globe (see appendix G.5). In order to validate the concept a pair of stereo images were acquired over the Witbank coal-mining area (F.3). Acquisition times differed by about one minute and thirty seconds.

Using a window width $w = 5$ and a neighbourhood size of 10, about 32.4 millions matches were retrieved. Their distribution is almost a Dirac centred on zero (see figure 4.22). This is compliant with the fact that the ortho-rectification was performed on a DEM retrieved from the panchromatic bands acquired at the same time as the multi-spectral imagery. The DEM is therefore up-to-date, hence leading to comparably few discrepancies in the mismatches.

The second maximum, located northwards, is likely to correspond to the discrepancies between the DEM and the actual digital surface model (DSM). In the DEM computation step, 'artificial' elevations such a houses or tree crowns are actually removed by means of the fast hybrid grey-scale reconstruction algorithm [Vincent, 1993; Arefi et al., 2011]. Ortho-rectification is therefore really performed using DEM, while imagery is actually acquiring image of surfaces. There is therefore potential error. Since the satellite is following a Sun-synchronous orbit, this is roughly flying north-south, the comparably high value at (0,-1) is simply accounting for this error.

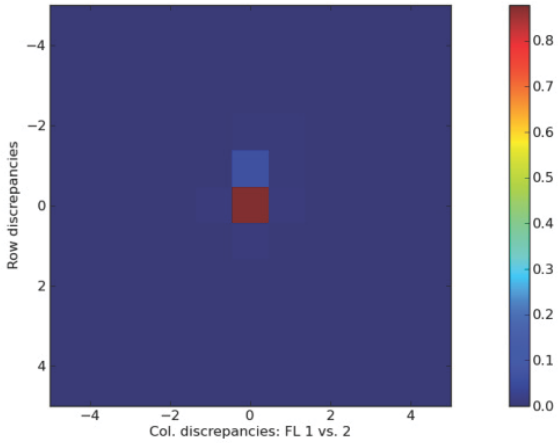

Figure 4.22: Histogram of geometric discrepancies for the Witbank 2011, World-View data-set (see F.3.1): $n = 10$, $w = 5$.

4.6 Conclusions

This chapter has presented an algorithm - CPAM - allowing the retrieval of pairs of pixels imaging roughly the same area from two different flight-lines. The method validation tends to indicate that at least 90% of the retrieved matches are correct, hence opening the door to further analysis.

The method was furthermore tested on several different data-sets, presenting different spatial, spectral and land-cover types. All tests led to dominantly successful matches, confirming that this algorithm is suitable for general use. Experiments have furthermore demonstrated its potential for dealing with satellite imagery. Besides the results presented here, the algorithm was also tested on multi-spectral thermal airborne data (AHS), as well as on other HyMap data-sets acquired over Sokolov in 2009.

In order to retrieve matches with a very small proportion of mismatches, pairs that are unsure are actually left away. This in turn potentially leads to some 'gaps' in the matches' coverage. Chapter 5 proposes some solution to this issues, while chapter 6 deals with the content extraction of the pairs retrieved by mean of CPAM.

Chapter 5

Filling the Gaps

For every credibility gap there is a gullibility fill.

Richard Clapton - Musician

Chapter Summary

This chapter provides methodologies for filling the 'gaps' present in the results from the previous section. A first part reviews existing methodologies and select a suitable one which are reviewed in more detail in a second part. Finally some practical computational issues are tackled and the resulting methodology is tested on practical examples.

5.1 Introduction

In section 4.1.2 the steps 3 and 4 of image registration, namely "*Transform Model Estimation*" and the "*Image Transformation*" steps - have been left aside: from a practical point of view, they correspond to an interpolation of the matched pixels over the unmatched ones. In the general use cases, this is likely to bring new uncertainties in the matching process, hence making further analysis quite complex, if not impossible. Such a process is indeed never completely error free [Simonson et al., 2007]. The introduced 'wrong matches' is in turn affect any further information extraction attempt.

Besides these pure interpolation errors, some pixels might actually have no matches at all. If in the first flight-line a certain pixel was representing a car, it is not impossible that the car did actually leave before the second flight-line occurred. The 'parking' pixel cannot be physically compared with the 'parking with car' pixel from the first flight-line. Performing an interpolation step however forces both pixels to form a pair. Similarly, over neat altitudes changes, some pixels might remain 'orphaned', having no equivalent in the second flight-line (e.g. see figure 5.1). Last but not least, due to fast platform pitch movements, some areas might be skipped in one flight-line, hence preventing any valid match retrieval over it. These three latter points are further arguing against an interpolation step when performing an analysis of the discrepancies.

On the other hand, it might occur that none, or too few, of the areas to be investigated are present in the retrieved pairs of pixels. This might for example be the case if one would like to include the effects induced by the orphaned pixels when analysing some classification results.

Section 5.2 reviews the existing methods to interpolate the pairs of retrieved pixels over the complete image. It should be noted that the terms interpolation is sometimes, for the sake

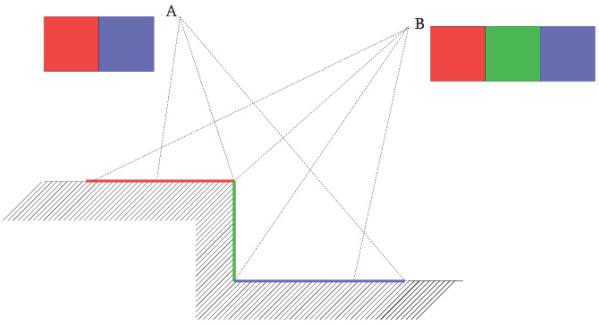

Figure 5.1: Example of orphaned pixel. This problematic scene imaged from viewpoint A will only produce two pixels while when imaged from B and additional green pixel will be retrieved.

of readability, substituted to the term approximation. When interpolating matches over the complete scene, original matches returned by CPAM are not modified. On the other hand, when dealing with an approximation, the original matches are used to provide a global approximation of the pairing process over the whole image. Once this has been done, they might have been slightly modified by the approximation process. This can typically be used to correct wrong matches.

5.2 Available Methodologies

Once a set of corresponding pixels have been retrieved, several methods can be used to extend these results over the complete overlapping parts. Given a set of N pixels from the first flight-line - $\{X_k = (x_k, y_k)\}$ - which got a match according to CPAM - $\left\{\tilde{X}_k = (\tilde{x}_k, \tilde{y}_k)\right\}$, the goal is to find a mapping F associating to each pixel from the first flight-line a match in the second flight-line under the constraint[1] $\forall X_k : F(X_k) = \tilde{X}_k$. It should be noted that, since the matches are not equally spaced only interpolation methods over irregular grids can be used.

The most widely used interpolation methods for image registration are probably the so-called global methods.

Given a pixel from the first flight line located at column x and row y, its coordinates in the second flight - $\tilde{x}$ and $\tilde{y}$ - line are provided by means of two P^{th} order polynomials with coefficients: $(a_{k,l})_{k,l}$ and $(b_{k,l})_{k,l}$: see equation 5.2.1. The polynomial coefficients are usually estimated by least square fitting on pairs of pixels retrieved by CPAM.

$$\begin{cases} \tilde{x} &= \sum_{k=0}^{P} \sum_{l=0}^{P} a_{k,l} x^k y^l \\ \\ \tilde{y} &= \sum_{k=0}^{P} \sum_{l=0}^{P} b_{k,l} x^k y^l \end{cases} \tag{5.2.1}$$

Extensions of this methodology include the projection model making use of first degree polynomials along with a first order normalisation factor, hence correcting basic perspective effects:

[1] The equal becomes a 'similar to' when performing an approximation instead of an interpolation.

$$\begin{cases} \tilde{x} &= \frac{a_0+a_1x+a_2y}{c_0+c_1x+c_2y} \\ \\ \tilde{y} &= \frac{b_0+b_1x+b_2y}{c_0+c_1x+c_2y} \end{cases} \tag{5.2.2}$$

Such global problems are, in general, not able to deal with local deformations such as the ones coming from outdated elevation information. They are therefore not suitable for this study's purposes. An alternative - at the cost of a greater computation intensity - consists of adaptating the polynomial coefficients to local areas. This can be done by performing a weighted least-square fitting when estimating the coefficients for a specific point, the weights decreasing with the distance to the investigated point [Goshtasby, 1988].

Another methodology for including local distortion in a global model consists in using radial basis functions or RBF, noted ϕ, as explained in equation 5.2.3. A radial basis function is a bi-variate application which is only depending on the distance between the input variables. Many different choices have been proposed such as Gaussian, splines or multi-quadrics. Radial basis function methods are known to provide very good results at the cost of a high computational complexity [Zitová and Flusser, 2003].

$$\begin{cases} \tilde{x} &= c_0 + c_1x + c_2y + \sum\limits_k w_k\phi(X, X_k) \\ \\ \tilde{y} &= d_0 + d_1x + d_2y + \sum\limits_k z_k\phi(X, X_k) \end{cases} \tag{5.2.3}$$

Another class of interpolation schemes are based on interpolation over triangulations. To start with, the pairs of retrieved control points are triangulated, for example getting a Delaunay triangulation. Akima [1978] proposed an interpolation based on fitting fifth order bi-variate polynomials inside each triangle, along with a smoothness constraint on edges. Goshtasby [1987] proposed a similar approach based on a cubic interpolation. This class of methods are usually designed to provide smooth solutions.

Another set of local methods are the kernel approximation ones [Myers, 1994]. Given an interpolation kernel, θ, the interpolation is provided by its convolution with found matches: see equation 5.2.4.

$$F(Y) = \sum \tilde{X} \star \theta(Y - X) \tag{5.2.4}$$

Silverman [1984] showed that when the kernel bandwidth is actually depending on the local density of initial points, this becomes an approximation of the spline interpolation. The main issue with this set of techniques is the lack of methodology to find a kernel suitable for the investigated data-set, which makes it comparably difficult to use in a generic context.

Finally kriging - also called Kolmogorov-Wiener prediction [Ligas and Kulczycki, 2010] - tries to interpolate point data based on a linear combination of the adjacent points. The related weights are then chosen in order to get a unbiased and with minimal variance estimation [Cressie, 1990]. It can be noticed that kriging and radial-basis functions interpolations are actually very similar and are actually equivalent in some specific cases [Myers, 1992, 1994].

Wiemker et al. [1996] compared global polynomial methods against Akima's method and the radial-basis function interpolation methods in the case of airborne imagery registration. Amongst different investigated radial basis function, the so-called thin-plate-splines, or TPS, lead to the bests results. Akima's method is in turn performing better than global polynomial methods while being faster than RBF interpolation. Since computational complexity is, in our case, a limited drawback, the radial basis function and more specifically the thin plate splines interpolation method are preferred.

The radial-basis methodologies is therefore preferred. Section 5.3 starts by explaining the method details, while section 5.4 provides the details about implementation.

5.3 Radial Basis Functions

The thin plate spline interpolation method was first introduced by Duchon [1977]. The method has since been used for image registration, especially in the context of medical imagery [Bookstein, 1989] and related topics such as haptic rendering [Höver et al., 2009], 3D-rendering [Carr et al., 2001] or even SONAR data processing [Barrodale et al., 1993]. The reason for such a broad scope of applications lies not only in their ability to cast both local and global effects, but also in their inherent design.

5.3.1 Principles

The thin plate spline interpolation method can be seen from a purely mechanical perspective. Given a thin plate of metal and a set of two and a half dimensional points, what would be the deformation of the plate allowing it to pass through each point while having a minimum bending energy?

Answering this question corresponds to solving the minimisation problem, where F is representing the plate surface:

$$\begin{gathered} \arg \min_{F:\mathbb{R}^2\to\mathbb{R}} \iint_{\mathbb{R}^2} \left|\frac{\partial^2 F}{\partial x^2}(x,y)\right|^2 + 2\left|\frac{\partial^2 F}{\partial x \partial y}(x,y)\right|^2 + \left|\frac{\partial^2 F}{\partial y^2}(x,y)\right|^2 dxdy \\ \text{such as} \\ \sum_{k=1}^{N} \left\| F(X_k) - \tilde{X}_k \right\| = 0 \end{gathered} \tag{5.3.1}$$

Duchon [1977] showed that solutions for this problem have to be of the form $\sum_{k=1}^{N} w_k \phi(X - X_k)$, where ϕ, are the so-called thin plate spline applied to the Euclidean distance between points and extended to 0 by continuity: see equation 5.3.2.

$$\begin{array}{rccl} & \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ \phi : & (X,Y) & \to & \begin{cases} \|X-Y\|^2 \ln \|X-Y\| & \text{if} \quad X \neq Y \\ 0 & \text{else} \end{cases} \end{array} \tag{5.3.2}$$

The N weights - $\{w_k\}$ - being virtually defined by the N constraints $F(X_k) = \tilde{X}_k$. So stated, the problem is however still ill-posed[2]. This can be tackled by adding a degree one polynomial of two variables along with the requirement that the weight coefficients vector is to be orthogonal to the polynomials of two variables of degree one, restricted to the set of control points [Buhmann, 2000; Duchon, 1977]. In other words, this corresponds to adding the requirement that polynomials of degree one of two variables - i.e. an affine transform - shall be preserved by interpolation.

The latter condition leads in turn to a further requirement. For the problem to be well-posed, it requires the control points to be non-colinear: this is in practice not an issue.

Thin plate splines are however not the only radial basis function that could be used. Hardy [1971] proposed to use multi-quadric functions: although not directly corresponding to a physical

[2]This does actually depend on the radial basis function used. In particular, the problem is well-posed when dealing with multi-quadrics [Buhmann, 2000].

interpretation, they are not requiring the polynomial orthogonality condition. Mitášová and Mitás [1993] introduced an 'anisotropic tension' term to the thin plate splines in order to take into account directional deformations. Buhmann [2000] presented several results valid on other of radial basis functions such as Gaussian kernels: $\phi_a(X,Y) = \exp -a\left(\|X-Y\|^2\right)$.

As stated, the TPS interpolation is fitting a surface on a set of control points: a slight arrangement has to be performed before being able to use it for interpolating the matched pairs.

The approach proposed by Bookstein [1989] for image registration consisted of considering it as two semi independent problems. A first interpolation considers $\tilde{x}\,(x,y)$ as a first 'surface', while a second one would obtain $\tilde{y}\,(x,y)$. Both results are then combined to obtain the final mapping: $\tilde{X} = (\tilde{x},\tilde{y})$. A corollary is that the 'bending' which is minimised is the bending of $\tilde{x}$ and $\tilde{y}$ as functions of (x,y).

It could be possible to use different RBF as the TPS ones. In the following, only the latter is investigated. The reason for this lies in their direct physical explanations and exhaustive documentation. This is furthermore in line with comparisons results made on remote sensing image co-registration experiments made by Wiemker et al. [1996], which indicated that TPS performed slightly better than multi-quadrics. Finally and although Mitášová and Mitás [1993] indicated that an anisotropy term might be profitable when dealing with terrain data, such a method requires setting the 'tension' parameter, which has to be done manually.

Stated as such, the interpolation problem becomes, equation 5.3.3, where the w_k and c_k (resp. z_k, d_k) are the unknowns. Section 5.3.2 reviews the practical solving of this interpolation problem.

$$\begin{cases} \tilde{x} &= c_0 + c_1 x + c_2 y + \sum\limits_k w_k \phi(X, X_k) \\ \tilde{y} &= d_0 + d_1 x + d_2 y + \sum\limits_k z_k \phi(X, X_k) \end{cases} \tag{5.3.3}$$

5.3.2 Solving the Interpolation

Given the requirement $\forall X_k : F(X_k) = \tilde{X}_k$ applied on equation 5.3.3, this yields N equations to be compared to the $N+3$ coefficients required. The three remaining equations are simply provided by the polynomial preservation condition on w_k (resp. z_k):

$$\begin{cases} 0 &= \sum\limits_{k=1}^{N} w_k \\ 0 &= \sum\limits_{k=1}^{N} w_k x_k \\ 0 &= \sum\limits_{k=1}^{N} w_k y_k \end{cases} \tag{5.3.4}$$

This can be directly expressed in terms of a classical linear system: $A \cdot Z = B$ where A, X and B can be subdivided as follow:

$$\begin{bmatrix} D & P \\ P^t & O \end{bmatrix} \cdot \begin{bmatrix} W \\ C \end{bmatrix} = \begin{bmatrix} L \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{5.3.5}$$

Where O is the three-by-three zero matrix, D a N by N matrix (equation 5.3.7) and P a N by 3 matrix as stated in equation 5.3.8.

$$O = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{5.3.6}$$

$$D = \begin{bmatrix} 0 & \phi(X_1, X_2) & \dots & \phi(X_1, X_N) \\ \phi(X_2, X_1) & 0 & \dots & \phi(X_2, X_N) \\ \vdots & \vdots & & \vdots \\ \phi(X_N, X_1) & \phi(X_N, X_2) & \dots & 0 \end{bmatrix} \tag{5.3.7}$$

$$P = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_N & y_N \end{bmatrix} \tag{5.3.8}$$

The weight vector, W simply contains the N w_k (resp. z_k) coefficients, and L the associated $\tilde{x}_k$ (resp. $\tilde{y}_k$). C is a three valued vector storing the c_k (resp. d_k). It should be noted that using this formulation, two inversions of A have to take place: one for computing the coefficients w_k and c_k associated with the columns - $\tilde{x}_k$ - and one more for obtaining their counterparts associated with the rows - $\tilde{y}_k$.

As stated from equations 5.3.5 to 5.3.2, the system is invertible, provided that the control points are not colinear. Franke [1982] however reported that the condition number of the matrix A is actually dependant on the points scaling and that an considerable improvement could be achieved when points are rescaled on the unit square. The interpolation would therefore be performed on the normalised positions: see equation 5.3.9.

$$\hat{X} = (\hat{x}, \hat{y}) = \left(\frac{x - \min\limits_k x_k}{\max\limits_k x_k - \min\limits_k x_k}, \frac{y - \min\limits_k y_k}{\max\limits_k y_k - \min\limits_k y_k} \right) \tag{5.3.9}$$

Once the interpolation has been performed, the points are being re-scaled to their origins by inverting 5.3.9. This trick allows for a reduction of the condition number of A by a factor 10^{12} to 10^{22} [Barrodale et al., 1993].

5.3.3 Dealing with Outliers

As seen in section 4.4, even if the vast majority of retrieved matches are valid ones, it is likely that some of the retrieved matches are actually outliers, without having a reliable way to identify them. Due to the continuity of the TPS solution, along with the condition $\forall k, F(X_k) = \tilde{X}_k$, this might lead to a blending of the issue to its neighbourhood pixels, hence amplifying the issues.

A partial workaround consists in slightly relaxing the interpolation condition, yielding an approximation instead of an interpolation [Rohr et al., 1996]. The minimisation is then based on equation 5.3.10, where λ is a positive regularisation parameter.

$$\begin{aligned} \arg\min_{F:\mathbb{R}^2 \to \mathbb{R}} \quad & \left(\lambda \iint_{\mathbb{R}^2} \left| \frac{\partial^2 F}{\partial x^2}(x,y) \right|^2 + 2 \left| \frac{\partial^2 F}{\partial x \partial y}(x,y) \right|^2 + \left| \frac{\partial^2 F}{\partial y^2}(x,y) \right|^2 dxdy \right. \\ & \left. + \sum_{k=1}^{N} \left\| F(X_k) - \tilde{X}_k \right\| \right) \end{aligned} \tag{5.3.10}$$

Itcan be seen that, when F is chosen in the form of an sum of thin plate splines as stated in equation 5.3.3, this corresponds to a generalisation of classical thin plate splines interpolation.

The bending energy and the choice of the internal form for F playing the role of prior information about its shape, while the sum of discrepancies can be seen as a data-term. When λ tends to zero, one comes back to the TPS interpolation. When λ tends to infinity, this corresponds to fitting only a polynomial of degree one to the data.

An interesting feature of this modelling corresponds to its ability to be solved by only introducing a slight modification of equation 5.3.5: this is setting the diagonal terms of D to λ instead of 0 [Wahba, 1990]. When λ is non-zero, this yields a slightly better conditioning of matrix A, by getting a non-zero diagonal.

The introduction of the approximation term however presents two major issues. To start with, the λ parameter has to be set manually. This typically involves several tries, depending on the data-set. The second drawback is slightly more problematic. Although the method filters the outliers out, it also smoothes the transitions in order to decrease the corresponding $\tilde{x}$ and $\tilde{y}$ bending energy (see figure 5.2).

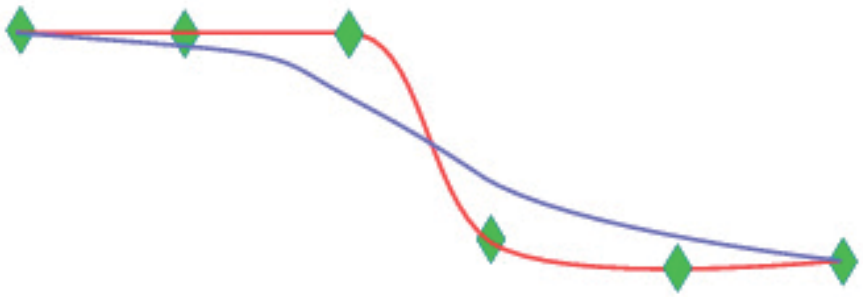

Figure 5.2: Analogy with a thin plate fitting with a regularisation parameter λ: the fitting becomes 'wrong' near brutal changes. Diamonds: control points; red line: interpolation curve; blue curve: approximated smoothed curve.

Both issues constitute an drawback of the method, thus limiting its application to the case where numerous erroneous outliers have been identified and where the actual solution is known to be very smooth.

5.3.4 Note on Complexity

Expressed as it is, the thin plate splines interpolation problem remains extremely computationally intensive and moreover requires a comparably high storage memory.

Just for storing A, in the general use case and taking into account its symmetry, $(3+N)(N+4)/2$ real values are required. When applied between flight-lines 3 and 4 of the Sokolov 2010 data-set, where about $5 \cdot 10^4$ pixels were retrieved, this yields a requirement[3] for 5 Gb just for storing A. When applied between flight-lines $4A$ and $4B$ of the Kaufbeuren 2012 test-site, N jumps to about $4 \cdot 10^6$ pixels, leading to a size requirement of 32 Tb which is not acceptable, even for a high-end desktop station.

Furthermore, besides being symmetric, A does not exhibit any interesting features for its inversion. This implies that a direct inversion has to be performed. This is typically achieved by means of a LU decomposition and therefore requires about $2N^3/3 + \mathcal{O}\left(N^2\right)$ floating point multiplications and divisions [Allaire, 2005; Buhmann, 2000].

[3] Using classical four-bytes floating point numbers.

Applied again to the flight-lines $4A$ and $4B$ of the Kaufbeuren scene, this reaches about $4 \cdot 10^{19}$ floating point operations. On a comparably high end processor being able to make 10^{11} floating point operations per second, this yields a computational time of at least 13 years. A speed up or an alternative therefore have to be examined.

Carr et al. [2001] pointed out that the use of radial basis functions with a compact support would produce a comparably sparse matrix, hence limiting both inversion and storage costs. The use of compactly supported RBF has however a drawback. Where there are 'large gaps' in the data, the interpolation process is not able to successfully bring information from border points over the gap centre, thus yielding comparably bad results over these areas. This is a noticeable drawback in the context of this work. As explained in chapter 4, CPAM is likely to perform poorly over homogeneous surfaces: this is in turn likely to result in large areas where few or no matches took place.

Another family of computational and storage speed-ups consists in discarding some of the input points. Iske [1999] for example proposed to remove points located in areas where the control points are dense relatively to the complete covered area. Höver et al. [2009] proposed three similar methods:

- a random selection of a subset of the control points.
- Creating a subset by iteratively adding the control points for which the current subset accounts the worst.
- Building an 'optimum' subset by means of bootstrapping.

These approaches are however based on data-sets where the control points are interchangeable. For this work, each control point however corresponds to a specific location and it makes little sense to remove one while keeping its neighbour.

Finally different methods based on model approximations have been proposed. Barrodale et al. [1993] are making use of the thin plate splines properties when applied on large distances to obtain a tabulation of the values which can then be interpolated. In [Powell, 1993], a similar approximation of linear combinations of TPS 'away' from control points was performed using a single truncated Laurent expansion. This is then further enhanced by transcribing the A into a $N-3$ square positive definite matrix, which can then be efficiently inverted by means of a preconditioned conjugate-gradient method. This leads in practice to a complexity of 'much less' than $\mathcal{O}\left(N^3\right)$ but does not tackle the storage issue.

The Beatson-Goodsell-Powell - or BGP - method [Buhmann, 2000] on the other hand is based on an iterative update of the coefficients. A proof of convergence has been made by Faul and Powell [1999]. The underlying idea consists in iteratively updating the thin plate spline coefficients based on local Lagrange functions. There is however no result on the BGP's speed of convergence. Experiments tend to indicate that in optimum cases, 1000 iterations are enough [Buhmann, 2000; Faul and Powell, 1999]. This approach however does not allow for an incorporation of the approximation parameter λ.

In the context of registration of remote sensing images, a much simpler workaround can however be used: image tiling. This is reviewed in section 5.4.

5.4 Retained Approach

In this specific case of registration, the notion of smoothness of the interpolation is to be understood in its *almost everywhere* sense. In other words, the elevation errors and the subsequent distortions are mostly continuous but faults can occur and are therefore tolerated. Moreover, the final results are pixels positions: this is discrete data, small intensity 'shocks' are therefore likely to get hidden when converting the results back to pixels positions.

Both considerations advocate for not requiring the complete interpolation to be 'smooth' but to relax this requirement into a 'local smoothness' requirement. This opens the way to a tiling of the interpolation process.

The retained method works as follows. The image is tiled in T blocks. For each block, the interpolation is computed using the M control points that are lying closest to its centre. When there are not enough control points within the tile, control points outside can be used, but the interpolation only occurs on the pixels located within the tile.

In order to obtain a quality assessment of the computations 10% of the control points, X_{k_l} are discarded from the computations and used as an external validation assessment. The corresponding validation metric being simply the root mean square error of the distances between retrieved matches - $\tilde{X}_{k_l}$ - and the interpolated matches: $F\left(X_{k_l}\right)$:

$$QA = \sqrt{\sum_{T} \sum_{l=1}^{M/10} \| \tilde{X}_{k_l} - F\left(X_{k_l}\right) \|^2} \tag{5.4.1}$$

In practice, M is set at a few hundreds of pixels. T is in turn be chosen in order to keep the same density of control points in the interpolation area as for a complete interpolation. If the complete overlapping part accounts S pixels, there are approximately S/T pixels per tile, which yields $T = NS/SM = N/M$. The complexity associated with the method is then reviewed in section 5.4.1.

5.4.1 Complexity

Just as in the previous section, N denotes the total number of matches as retrieved by CPAM, the number of control points used for each interpolation is noted M, while the total number of tiles to be investigated is noted T and the total number of pixels in the original image is S. This yields approximately S/T pixels per tile.

Each interpolation, as before, yields a computational complexity in $\mathcal{O}\left(M^3\right)$, the global complexity is therefore in the magnitude of $\mathcal{O}\left(TM^3\right)$ to be compared with the original $\mathcal{O}\left(N^3\right)$. When adding the requirement that the tiled interpolation has to be based on the same control points density as the global one, one gets the relationship $N/S = \mathcal{O}\left(MT/S\right)$, i.e. $N = \mathcal{O}\left(MT\right)$. This yields an improvement of a factor T^2, up to the scaling constants.

A further interest of the tiling approach consists in its direct transcription into parallel code: each tile can indeed be independently processed. When C computational threads are available, this yields a final complexity of $\mathcal{O}\left(TM^3/C\right)$.

In terms of memory, the gains are also significant: each tile required $\mathcal{O}\left(M^2\right)$ floating point. With M being typically about a few hundreds pixels, this space requirement becomes largely acceptable.

5.4.2 Application

This section presents some illustration of the interpolation method to some exemplary data-sets.

Sokolov 2010

As an illustration, the interpolation process was run between flight-lines 2 and 3 of the Sokolov 2010 data-set. As first run, with $\lambda = 0$ - pure interpolation - $M = 100$ and $T \approx 160000$, produced a root mean square error for the validation pixels of 0.0013 pixel. A subset of the interpolation results is provided on figure 5.3. Computation time was of 8 seconds, without parallelism.

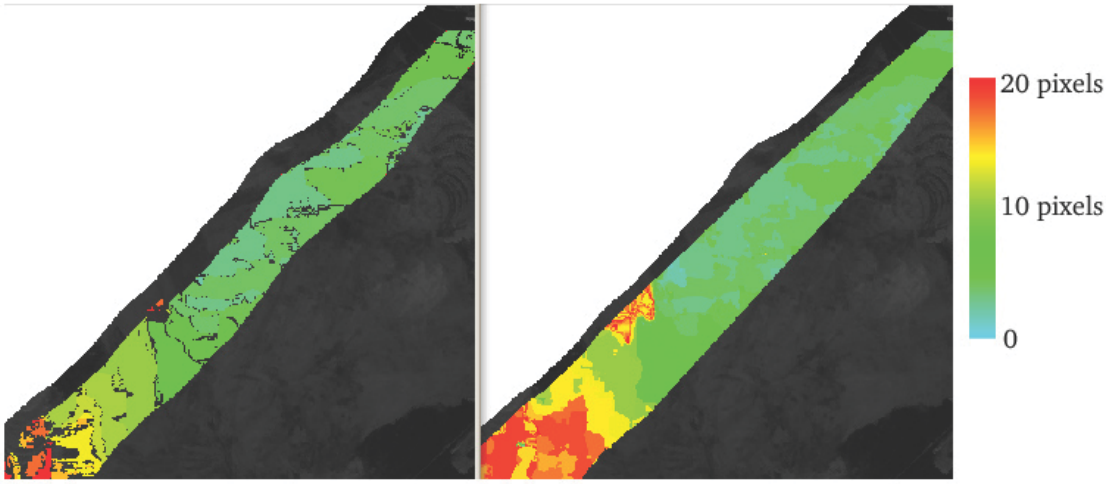

Figure 5.3: Exemplary interpolation between flight-lines 2 and 3 of the Sokolov 2010 data-set with $\lambda = 0$. Left: before interpolation, right: after interpolation. Coloured pixels represent matches, their hue depending on the observed discrepancies.

A close investigation of the results reveals that, here and there, errors are occurring on the tiles' edges. The effect is however comparably rare and are therefore neglected. Another slight issue are the 'noisy' interpolation results over areas where, besides outliers, only few matches were found. The interpolation altogether does however seem reliable.

Kaufbeuren

The method was finally used to solve the interpolation issue on the Kaufbeuren 2012 data-set between flight-lines $4A$ and $4B$. The validation pixels provided an estimation of the interpolation root mean square error of 0.0008 pixels for $M = 100$, $T \approx 67000$ and $\lambda = 10$. A subset of the results is presented on figure 5.4: notice the low amount of matches over the forest, a homogeneous area - likely to be false ones. Here the use of a smoothing parameter helps limiting the impact of outliers. Again, the interpolation does not lead to any obvious issue, while the validation root mean square error score is very low.

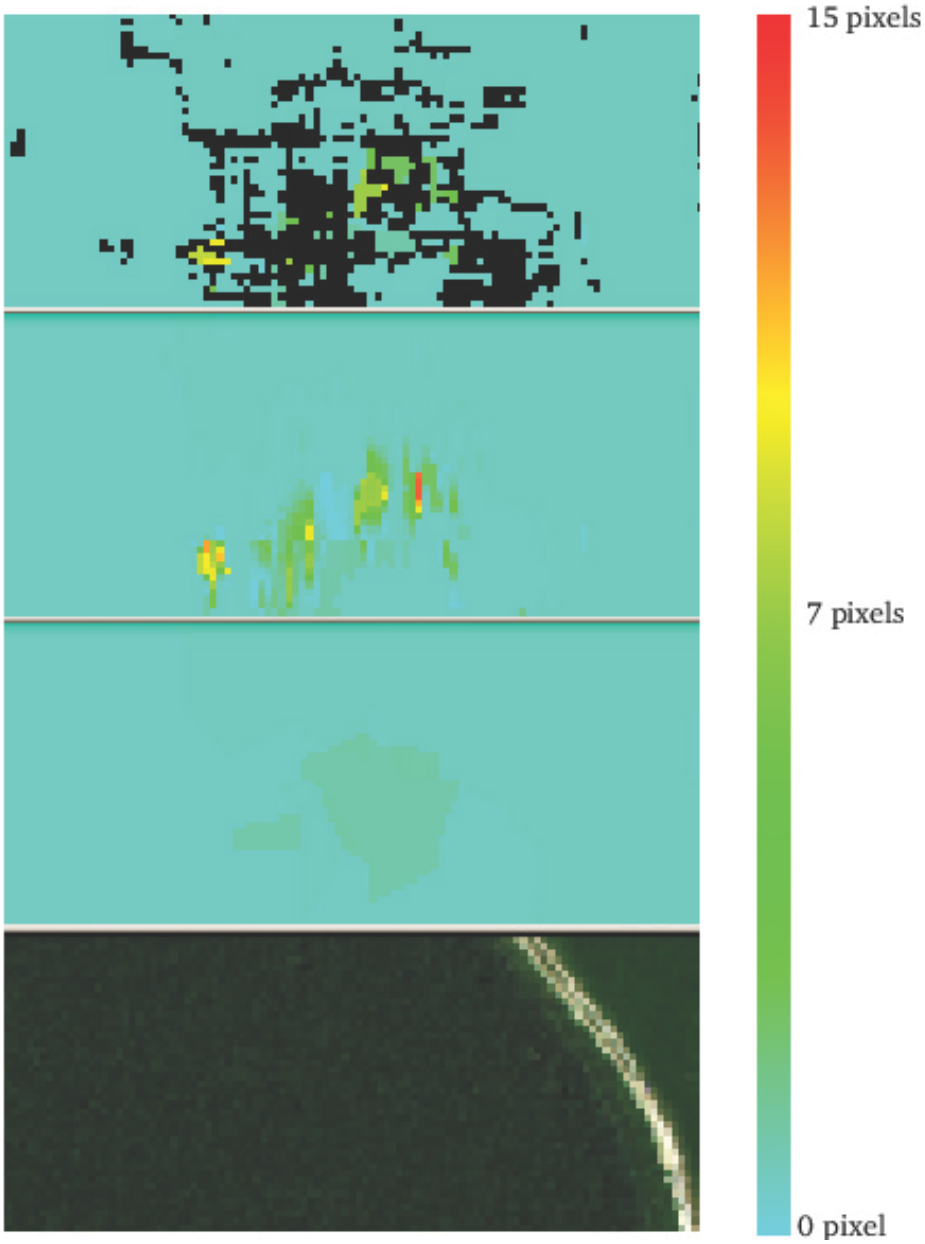

Figure 5.4: Exemplary interpolation result between flight-lines $4A$ and $4B$ of the Kaufbeuren 2012 data-set with $\lambda = 10$. Hue represent matches discrepancies between theoretical match location, and actual match locations. Top: original matches, second: interpolated image - no smoothing, third: interpolation with smoothing, bottom, false colours imagery.

5.5 Conclusions

A method to interpolate the retrieved matches over the complete overlapping parts of the imagery has been proposed. It can be operationally run on the different data-sets without inducing too wide discrepancies in the matches.

It should however be stressed again that pairs of pixels retrieved after an interpolation step should not be used for performing any in-depths analysis of the radiometric discrepancies. There is for example absolutely no physical reason to analyse discrepancies between an orphaned pixel as presented in figure 5.1 and one of its neighbours.

Interpolation of results should therefore be strictly restricted to the following uses:

- global or semi global discrepancy description, for example for providing a statement of classification discrepancy percentages between both flight-lines. In such cases the final user is likely to be interested in the overall changes of classification between both flight-lines, including discrepancies due to the presence of orphaned pixels.

- the area of interest has a too poor coverage in terms of matched pairs of pixels. In such cases, the area is however likely to be a homogeneous area and some other form of quality assessment might make more sense.

Chapter 6

Quality Indicators for Maps

He had bought a large map representing the sea,
Without the least vestige of land:
And the crew were much pleased when they found it to be
A map they could all understand.

"What's the good of Mercator's North Poles and Equators,
Tropics, Zones, and Meridian Lines?"
So the Bellman would cry: and the crew would reply
"They are merely conventional signs!"

Lewis Caroll - Mathematician

Chapter Summary

This chapter presents how results derived from CPAM can be used to provide a quality assessment in terms of reliability. A first part reviews direct assessment techniques as applied to CPAM outputs. In a second phase, a critical review of existing higher-level methods takes place, focusing especially on the well-known Cohen's κ. Finally, the proposition is made to bring Krippendorff's α for evaluating remote sensing data. These notions are illustrated with examples.

6.1 Introduction

Chapter 4 has explained how to retrieve pairs of observation of the same geo-location. This chapter in turn explains how these pairs of information can be used to obtain information about the quality of so-called land-products - or L3 - products.

A direct approach consists of comparing the mapping application on both flight-lines. This is: comparing results obtained for pairs of matched pixels. The further apart they are, the worse the land-product is. Since this comparison takes place between the overlapped parts of flight-lines, they can be performed on virtually every scene containing more than one flight-line. Moreover, since comparisons are made over flight-lines fringes this comparison is made in a 'worst case' variability: the BRDF effects are indeed maximised.

This comparison corresponds, in the generic use case, to a repeatability assessment: parameters such as the viewpoint are changing between two flight-lines. In some cases however, when two flight-lines are imaged with the same viewpoint at almost the same time, this can be extended

to a repeatability assessment technique. The direct link with the Guide to the Expression of uncertainty however stops here. This assessment is definitely not a type B uncertainty evaluation, while on the other hand, it can be seen as a type A assessment of uncertainty associated with the discrepancies between two flight-lines, not of the mapping uncertainty itself.

On the other hand, such a comparison has two inherent drawbacks. To start with, results obtained by this means are scene dependant. Comparing discrepancies between different mapping results implies sharing similar land-covers and reflectance. In practice, that almost never occurs due to varying atmosphere properties, soil water content, vegetation health, etc. This drawback is however overcome by the method ability to be run on both scenes, thus rending null and void the need to extrapolate the results from one scene to another.

The second drawback consists of the difficulty to link discrepancies back to their origins. In other words, it is only stating 'how good' the mapping is without explicitly stating what bottle necks are preventing improvements. This last issue is partially tackled in section 7. Finally, it should be noted that this cross comparison can only account for errors, bias - being by definition present in both flight-lines - remaining hidden when comparing both results.

The rest of this chapter is organised as follows. Section 6.2 presents basic methodologies along with some examples. Section 6.3 provides an overview of advanced methodologies before applying them to illustrative mapping results in section 6.5.

6.1.1 Principles

An acceptable quality?

Before going further, one might ask what an acceptable quality - i.e., in this study, reproducibility - has to be met. Smits et al. [1999], quoting [Lins, 1994], reported quite high discrepancies in terms of users requirement: see table 6.1. This can easily be understood by comparing different applications. Imagine the following cases:

- It has been reported in public media that remote sensing could be used to map swimming-pools in order to target tax-audits on potential tax-dodgers [Steinvorth, 2010]. One can argue that when this mapping is only slightly better than random audits, the objective is met. Since pure binary random checks correspond to a reproducibility of 25%, a reproducibility of only 50% would be quite acceptable.

- Lets consider now an early warning system for volcano eruptions. An erroneous alert and the subsequent evacuations might have drastic economical consequences. On the other hand, failing to identify an upcoming eruption might lead to fatalities. The corresponding accuracy and reproducibility requirements therefore has to be extremely high. As a comparison, and since there is no standardised guideline defined for remote sensing, the requirements defined for civilian aviation is used. The FAA [2014] defined an average per flight risk of 10^{-9} for level 4 motorisation events - i.e. events inducing forced landing or loss of an aircraft or serious injuries or fatalities[1]. Meeting this requirement does in turn imply a functional reproducibility being at least better of 99.99999999%...

These two examples stress the impossibility to provide a generic reproducibility assessment threshold value. One could of course argue that 'the more, the better'. In practice, granting a better reproducibility would however induce other constraints such as greater costs [Smits et al., 1999]: getting the optimum quality is therefore a trade-off between users' need and technical feasibility and available funding. The 'right' suitable quality is therefore application dependent

[1] This rate corresponds to the probability of an average human being dying within the next hour.

Accuracy Requirement	Percentage of Users
95%	31%
90%	57%
85%	10%
80%	3%
75%	1%

Table 6.1: Differences in terms of quality requirements - in terms of classification accuracy - made among 405 remote sensing data users: [Smits et al., 1999], citing [Lins, 1994].

and the following only focuses on stating the reproducibility, without trying to access whether it is suitable for a specific purpose or not [Krippendorff, 2004]. The results are therefore only linked to a 'standard' accuracy requirement for remote sensing data - 85% - even if this has no practical meaning [Smits et al., 1999]. Such a value indeed depends on the actual end-user needs, while the term 'accuracy' has to be linked with a specific quantity.

Data types

At this point, one last issue has to be tackled: the data type returned by the mapping. The type of data - or level of measurement - provided by the mapping application influences the comparisons. Comparing two classification types cannot be performed in the same way as when comparing two numerical values. This section quickly presents the different data types along with suitable metrics.

In general, several mapping data types - or measurement scale types - can be identified. From the initial four types proposed by Stevens [1946], Chrisman [1998] derived an extended set of ten types suitable for geographical data. Krippendorff [2011a] implicitly proposed metrics for six of them. Three of the remaining led to straight forward extensions of the exiting metrics to their specific cases. The last one, the *derived ratio* data type, corresponding to a product of ratios of fundamental properties, is set aside since its is only discriminated in terms of its origins, not in terms of the data it contains.

- Nominal data [Stevens, 1946]: this corresponds to the case where data is actually only a label. This is typically the type of data associated with a classification result. A straightforward metric for the relationship of two classes, k and l is the Kronecker function (see equation 6.1.1 [Krippendorff, 2011a]).

$$\mu_n(k,l) = 1 - \delta_{k,l} = \begin{cases} 0 & \text{if} \quad k = l \\ 1 & \text{if} \quad k \neq l \end{cases} \tag{6.1.1}$$

- Graded membership data [Chrisman, 1998]: this correspond to a sort of 'soft' ordinal data. Data corresponds to a nominal data along with a degree of membership to the label prototype. This is for example the case when dealing with a classification where not only the class but also the distance to the class description are available. In spectroscopy, this could for example be the spectral-angle between the class (e.g. the end-member spectrum) and the investigated pixel. Noting $d(k, c_k)$ the corresponding distance between element k and its class - c_k - a straightforward metric can be seen as a weighted Kronecker function:

$$\mu_g(k,l) = \frac{(d(k,c_l) + d(l,c_k)) - (d(k,c_k) + d(l,c_l))}{2 \max\limits_{k,l} d(c_k, c_l)} \tag{6.1.2}$$

One can see from equation 6.1.2 that when the classification is logical - i.e. $\forall (k,l), d(k,c_k) \leqslant d(k,c_l)$, equation 6.1.2 always yields positive results.

- Ordinal data [Stevens, 1946]: this type is used when data is encoded as an ordering, i.e. each element gets a rank, but where no notion of 'degree of difference' can be defined. In remote sensing, this might be used to represent qualitative data (e.g. Low, Medium, High). Noting n_r the number of elements associated with rank r, a suitable metric for evaluating discrepancies between the k^{th} and l^{th} ranks is given in equation 6.1.3, max and min being the extremal ranks [Krippendorff, 2011a].

$$\mu_o(k,l) = \sqrt{\sum_{r=k}^{l} \left(n_r - \frac{n_k + n_l}{2} \right)^2} \tag{6.1.3}$$

- Interval data [Stevens, 1946]: this type of data can be used to retrieve a degree of difference between two values but not a ratio. It is for example associated with Celsius temperatures: since the definition of the zero is arbitrary, a ratio of two temperatures would not have any physical meaning: $20C$ is not twice as hot as $10C$. A suitable metric between two values is the 'offset' between two results (see equation 6.1.4 [Krippendorff, 2011a]).

$$\mu_i(k,l) = l - k \tag{6.1.4}$$

- Log-Interval data [Chrisman, 1998]: the log-interval data is the counterpart of interval data when a ratio but no difference can be retrieved: its classical use corresponds to earthquake intensity measurement. A straightforward metric can be derived from equation 6.1.4 by playing on the logarithms:

$$\mu_l(k,l) = \log(l) - \log(k) = \log(l/k) \tag{6.1.5}$$

  Note that since ratios can be defined, k and l are necessary non zero. Since the difference degree is not defined either, they must by sharing the same sign, the logarithm of their ratio is therefore defined.

- Ratio data [Stevens, 1946]: this is an extension of the interval data when an absolute zero is defined but no absolute scaling unit can be retrieved. Chrisman [1998] defines it as the *extensive ratio* data. This is typically the data type associated with distances, time frames, etc... A suitable metric between two values is the 'normalised distance' extended to 0 when, $k = l = 0$: see equation 6.1.6. Note the analogy with NDVI-like indexes which are de facto scaling independent.

$$\mu_r(k,l) = \frac{k-l}{k+l} \tag{6.1.6}$$

- Periodic data [Chrisman, 1998]: this corresponds to data exhibiting periodicity and is typically associated with angles. The discrepancy between two values k and l can then be defined in terms of the shortest 'distance' between them given the period T: see equation 6.1.7 [Krippendorff, 2011a].

$$\mu_p(k,l) = \sin \frac{\pi(k-l)}{T} \tag{6.1.7}$$

- Counts [Chrisman, 1998]: These correspond to numbers of elements. A zero value as well as a scaling unit are therefore clearly defined but no ratio can be defined. A straightforward metric is obtained by the difference count:

$$\mu_c(k,l) = l - k \tag{6.1.8}$$

- Absolute data [Chrisman, 1998]: this is an extension of the ratio data where the definition domain is bounded, hence setting the unit (scale) in an absolute way. Probabilities and reflectance are typically of the latter type, since both have to be laid in the interval $[0;1]$. A suitable metric can be defined by means of the higher and lower bound, respectively b_{max} and b_{min}: see equation 6.1.9 [Krippendorff, 2011a].

$$\mu_a(k,l) = \begin{cases} 0 & \text{if} \quad k = l = b_{min} \text{ or } k = l = b_{max} \\ \dfrac{k-l}{\sqrt{(k+l-2b_{min})(2b_{max}-k-l)}} & \text{else} \end{cases} \tag{6.1.9}$$

It should be pointed out that these distinctions might become limited when dealing with a complete geographical data-set, especially when dealing with relationships instead of properties [Chrisman, 1995, 1998]. In this work, the focus is however set on 'basic' $L3$ products expressed in form of rasterised imagery, where each pixel is indeed defined by its properties. There is therefore no need for any more complex framework than the presented one.

Finally, these measurement levels are only studied from their ability to describe different types of data, not from a measurement theory perspective: this is the reason why the 'derived ratio' type is ignored. One could for example easily argue about the work presented by Chrisman [1998] by stating that from his perspective of classical mechanics, a energy is a ratio data. However from a relativistic or quantistic point of view most of these considerations just falls apart: indeed $E = mc^2$ implies the energy to be a derived ratio. On the other hand, quantification of the energy levels in an atom would plead for it to be an ordinal data.

Chrisman [1998] moreover stated that the proposed categories were '*by no mean complete*': experience shows that the huge majority of mapping applications lead to one of these categories and that there is no need for extending them.

6.1.2 Applications

The concepts presented in this chapter were applied to several case studies. Although they are usually straightforward to describe, the test case for nominal data is requiring some explanations. It corresponds to a classification based on the Sokolov 2010 data-set (see section F.2.1). The mapping application was a step wise spectral angle mapper based classification, performed by Dr. Rogge, based on 59 end-members. These were extracted by means of the Spectra-Spatial-End-members Extraction (SSEE) algorithm [Rogge et al., 2007]: the list of end-members is provided in table 6.2. It should be stressed that this is by no means an actual 'final' classification. In particular no post-classification step has been performed, thus leaving room for improvements. Furthermore the classes are not homogeneous as they actually are a mixture of feature-classes (e.g. deep clay 1), land cover classes (e.g. soil) and functional classes (e.g. Urban classes). This apparent roughness does nevertheless allow for a straightforward interpretation of the classification results and is therefore used for applicative examples.

Class Index	Description	Class Index	Description
(1)	(no data)	2	Unclassified
3	Coal type 1	4	Coal type 2
5	Coal with clay	6	Iron oxyde
7	Iron narrow clay	8	Deep clay type 1
9	Deep clay type 2	10	Iron broad clay type 1
11	Iron broad clay type 2	12	Kaolinite type 1
13	Kaolinite type 2	14	Gypsum
15	Soil	16	Brown vegetation type 1
17	Brown vegetation type 2	18	Water with vegetation
19	Vegetation type 1	20	Vegetation type 2
21	Vegetation type 3	22	Vegetation type 4
23	Vegetation type 5	24	Vegetation type 6
25	Vegetation type 7	26	Dark water
27	Green water	28	Red water
29	VNIR-features 1	30	VNIR-features 2
31	VNIR-features 3	32	VNIR-features 4
33	Urban type 1	34	Urban type 2
35	Urban type 3	36	Urban type 4
37	Urban type 5	38	Urban type 6
39	Urban type 7	40	Urban type 8
41	Urban type 9	42	Urban type 10
43	Urban type 11	44	Urban type 12
45	Urban type 13	46	Urban type 14
47	Urban type 15	48	Urban type 16
49	Urban type 17	50	Urban type 18
51	Urban type 19	52	Urban type 20
53	Urban type 21	54	Urban type 22
55	Urban type 23	56	Urban type 24
57	Urban type 25	58	Urban type 26
59	Urban type 27	60	Urban type 28
61	Urban type 29		

Table 6.2: Description of end-member classes obtained by SSEE [Rogge et al., 2007] over the Sokolov 2010 test site.

6.2 Basic Analysis of L3 products

This section reviews some crude but simple to use reproducibility metrics that could be used to compare matches as retrieved by CPAM. Section 6.2.1 reviews some generic tools along with means to derive corresponding confidence intervals in section 6.2.2. Section 6.2.3 extends them to some specific data types. Finally sections 6.2.4 and 6.2.5 presents some applications for these concepts.

6.2.1 Generic Analysis

In the most generic case, when N pairs of matched pixels $\left\{\left(X_k, \tilde{X}_k\right)\right\}_k$ have been retrieved for a scene S, and when a mapping application, $M : X \rightarrow M(X)$, was run on each flight-line, a basic generic reproducibility assessment is provided by the average discrepancy: equation 6.2.1.

$$AD(M,S) = \frac{1}{N}\sqrt{\sum_{k=1}^{N} \mu_M\left(M(X_k), M\left(\tilde{X}_k\right)\right)^2} \tag{6.2.1}$$

Where μ_M is a metric suitable for evaluating discrepancies relative to the mapping results. For example, when the data is interval data, AD becomes the classical root mean square error equation.

A second quality metric of interest is the maximum discrepancy, MD, as expressed in equation 6.2.2: it provides the greatest discrepancy obtained by running M on scene S. Although it is strictly speaking accounting for the notion of worst case scenario, it might be of lesser interest when dealing with nominal data. In such a case, the value can only be zero or one:

$$MD(M,S) = \max_k \left|\mu_M\left(M(X_k), M\left(\tilde{X}_k\right)\right)\right| \tag{6.2.2}$$

This leads to the following question: in the general case, what information is actually brought by a given value of AD. Although 0 is clearly indicating a perfect reproducibility, higher values seem to be more difficult to assess, except for the usual statement 'the greater, the worse'. This issue is addressed in further details in chapter 7.

Slightly more advanced but more applicable quality indicators is the cumulative density distributions of the discrepancies. The cumulative function accounts for the proportions of discrepancies lower or equal to a certain threshold ϵ: it is therefore an increasing function taking its values in $[0;1]$.

$$CD : \begin{cases} [0;+\infty[& \rightarrow & [0;1] \\ \epsilon & \rightarrow & \dfrac{\sum\limits_{k|\mu\left(X_k,\tilde{X}_k\right)\leqslant\epsilon} \mu\left(X_k,\tilde{X}_k\right)}{\sum\limits_{k=1}^{N} \mu\left(X_k,\tilde{X}_k\right)} \end{cases} \tag{6.2.3}$$

A straightforward application of the cumulative distribution is to obtain a link between acceptable errors and reproducibility. For example, when one of the mapping requirement was to grant an reproducibility greater than R_0. What minimal error ϵ on the results can to be tolerated to meet this requirement? An inversion of CD leads directly to an estimate of the minimal acceptable error (see figure 6.1). Similarly when given a requirement on the maximal admissible error ϵ_0, the achieved reproducibility is directly given by $CD(\epsilon_0)$.

At last, one can similarly define the discrepancies density function accounting for the proportion of matches associated with a specific interval of discrepancies:

$$DD : \begin{cases} [0;+\infty[& \rightarrow & [0;1] \\ \epsilon & \rightarrow & \dfrac{\sum\limits_{k|\mu\left(X_k,\tilde{X}_k\right)\leqslant\epsilon} \mu\left(X_k,\tilde{X}_k\right)}{\sum\limits_{k=1}^{N} \mu\left(X_k,\tilde{X}_k\right)} \end{cases} \tag{6.2.4}$$

It can finally be noted that a by-product of these computations is a discrepancy localisation map: since each pair of pixels is associated with both a localisation in each flight line and a mapping discrepancy, a 'discrepancy map' can be produced (e.g. figure 6.2).

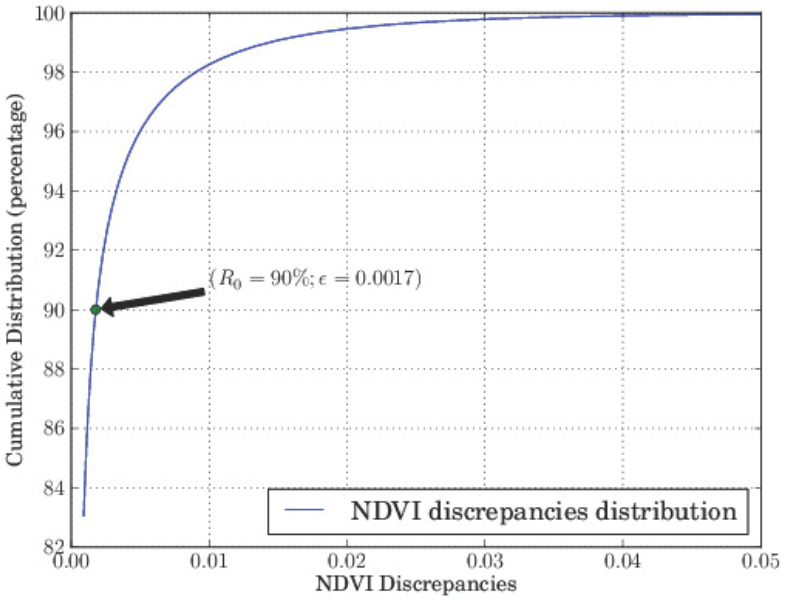

Figure 6.1: Example of cumulative distribution: absolute data - NDVI results over the WorldView II scene over Witbank. $R_0 = 90\% \Leftrightarrow \epsilon_0 = 0.0017$. $MD = 0.37$, $AD = 5.33e-06$.

6.2.2 Confidence Intervals

Chapter 3 has already indicated that several entities such as QA4EO or the GUM have stressed the need to provide confidence intervals. In general, this can not always be performed on the L3 product itself. For example, in the specific case of nominal data, a confidence interval has little meaning when applied to the mapping itself. On the other hand, since AD and MD are both ratio data, one can retrieve their own confidence intervals which are linked with the data.

Still in general, little if anything is known about the distribution of AS and MD. There is therefore no analytical expression of the corresponding confidence intervals and one has to rely on non parametric bootstrapping [Hesterberg et al., 2007; Janssen and Pauls, 2003] as explained in the next paragraphs.

Principles

When attempting to characterise a property from a population, one usually proceeds by taking a sample and computing the property value based on this population subset. A direct validation of this result might be obtained by taking a second sample, computing the property value for this second subset before comparing both results. In other words, inferring a property from a population sample should be taking into account the a priori random variation of this sample. In our case the property to be measured - AD or MD - would all the same be varying depending on retrieved matches and this variability has to be estimated too. In order to assess this variability, a straightforward method would consist of looking at what is happening when only a subset of the retrieved matches were used to compute the parameter of interest. This can be achieved as

Figure 6.2: Example of a discrepancy map: case of classification (nominal data) over Sokolov 2010. Green pixels: no observed discrepancy - red pixels: classification discrepancy - white pixels: no match or no data.

stated in algorithm 3 [Hesterberg et al., 2007].

Algorithm 3: Bootstrapping

Data: A spectral metric: $\mathcal{S}(\bullet, \bullet)$
Data: N pairs of matched pixels, $\left(X_k, \tilde{X}_k\right)$
Data: θ a parameter whose statistics should be computed: AD or MD
Data: B a number of iterations
Result: L the list of retrieved pairs

Initialisation:
Compute θ for the complete data set: θ_0
for b *from* 1 *to* B **do**
    Select randomly N pairs with replacing
    Compute and store θ for this set: $\tilde{\theta}(b)$
end
`// Compute the bootstrap mean:`
$\bar{\theta} = 1/B \sum_{b=1}^{B} \tilde{\theta}(b)$
`// Compute the bootstrap standard error:`
$\sigma_\theta = \sqrt{\sum_{b=1}^{B} \left(\tilde{\theta} - \bar{\theta}(b)\right)^2 / (B-1)}$
`// Compute the bootstrap cumulative distribution:`
$G := \theta \rightarrow Card\left\{\tilde{\theta}(b) \,|\tilde{\theta}(b) < \theta\right\}/B$
`// Compute the bootstrap density function:`
$g := \theta \rightarrow \dfrac{\partial G}{\partial \theta}(\theta)$

One can note from algorithm 3 that G is actually not derivable: g has therefore to be understood - in practice - as a limit case, i.e. as being the $\left\{\tilde{\theta}(b)\right\}$ histogram. Similarly G is often represented as a linear or smooth interpolation between the different values taken by $\tilde{\theta}(b)$. These mathematical nuances however play little role in the following.

It should also be noted that the bootstrap mean and standard error can be very different from the original population mean and standard deviation. In order to illustrate this behaviour, let imagine a sample presenting a huge outlier whose value is affecting both the population mean and standard deviation. If bootstrapped, this specific sample is unlikely to be present in each re-sampled data-set and does therefore not contribute uniformly to $\bar{\theta}$ and σ_θ.

This is the reason why, once the bootstrap has been performed, one can define the bootstrap bias as being $\bar{\theta} - \theta_0$. A large bias indicates that the obtained value for θ_0 is actually not representative for the expectancy for θ distribution within the scene.

Finally a last point has to be cleared: what is the suitable value for the number of re-sample to be evaluated - B. Most sources - [Carpenter and Bithell, 2000; Efron, 1984; Hesterberg et al., 2007] recommend to take $B \approx 2000$. Other authors working on related bootstrapping methods - e.g. [Davidson and MacKinnon, 2001; Pattengale et al., 2010] - did however point out that B does actually depend on other factors. Krippendorff [2006] uses for similar bootstrapping application $B = 20000$. From a theoretical taking N samples among N with replacing yields a large total of possibilities:

$$\left(\begin{array}{c} 2N-1 \\ N \end{array} \right) = \frac{(2N-1)!}{N!(N-1)!} \tag{6.2.5}$$

This number quickly explode: for N equals 20, this yields about 7 billion possibilities. A random sub-sample of only a few thousands seems therefore unsuitable for a practical case where N is about a few millions - as in most results from CPAM. The number of re-sample to be performed is therefore adapted from application to application. For most of them, values of B up to 100000 have proved to be suitable.

Towards Confidence Intervals

When the bias is small, a straightforward way of getting of the $P\%$ confidence interval corresponds to equation 6.2.6 where $\mathcal{Q}_x(\mathcal{B}_d)$ is the $x\%$ centile of the Bootstrapped distribution $\mathcal{B}_d$ [Hesterberg et al., 2007].

$$\left[\mathcal{Q}_{50-P/2}(\mathcal{B}_d); \mathcal{Q}_{50+P/2}(\mathcal{B}_d)\right] \tag{6.2.6}$$

When the bootstrapping bias is not negligible, one has to rely on slightly more complex methods. The reference methods [Hesterberg et al., 2007] are namely the 'Bias-Corrected accelerated' BC_a being slightly slower but producing more accurate results and the Bootstrap tilting interval (TIL) producing slightly worse results but being faster to run [DiCiccio and Efron, 1996]. [Carpenter and Bithell, 2000] analysed available methodologies - including BC_a and TIL and concluded that BC_a is - in all use cases - a very suitable method, while TIL might perform poorly in some specific cases. Since computation time is here not a strict requirement, BC_a is preferred.

The principle of BC_a consists in correcting the confidence interval as expressed in equation 6.2.6 in order to account for both bias and skewness by means of a so-called bias correction - z_0 - and an acceleration a. Once these two parameters have been computed, the $P\%$ confidence interval is given by $\left[\tilde{\theta}_{BC_a}(0.5 - P/200); \tilde{\theta}_{BC_a}(0.5 + P/200)\right]$ where $\tilde{\theta}_{BC_a}$ is defined in 6.2.7, Φ being the standard normal cumulative density function.

$$\tilde{\theta}_{BC_a} \begin{cases} [0;1] & \to \mathbb{R} \\ z & \to \quad G^{-1}\left[\Phi\left(z_0 + \dfrac{z_0 + \Phi^{-1}(z)}{1 - a\left(z_0 + \Phi^{-1}(z)\right)}\right)\right] \end{cases} \tag{6.2.7}$$

One can note that when both z_0 and a are zero, equation 6.2.7 yields back to 6.2.6. The bias correction parameter is simply retrieved as in equation 6.2.8.

$$z_0 = \Phi^{-1}\left[\frac{Card\left\{\tilde{\theta}(b) < \theta_0\right\}}{B}\right] = \Phi^{-1}\left[G(\theta_0)\right] \tag{6.2.8}$$

The acceleration coefficient is slightly more complicated to obtain. In its simplest form[2], it can be computed by means of equation 6.2.9 [DiCiccio and Efron, 1996].

$$a = \frac{1}{6}\frac{\sum_{k=1}^{N} U_k^3}{\left(\sum_{k=1}^{N} U_k^2\right)^{3/2}} \tag{6.2.9}$$

Where the U_k are the so-called jackknife influence functions as stated in equation 6.2.10: θ_0^k being the value obtained for θ_0 when the k^{th} pair is discarded from computations.

$$U_k = (N-1)\left(\theta_0 - \theta_0^k\right) \tag{6.2.10}$$

The coverage error when using this method is extremely small: $\mathcal{O}(1/N)$, hence making BC_a a very efficient method [Carpenter and Bithell, 2000], especially for this work, where N is typically very high.

6.2.3 Specialised 'Basic' Analysis

Several generic tools have been presented for analysis L3 products in the general case. However, even if they can be applicable to each type of data, the possibilities offered by their results analysis remain quite basic. The following sections therefore review some more specific but also more adapted tools.

Nominal Data: Relative Cross-Confusion Matrices

When dealing with nominal data, typically the outcome of a classification, several analysis tools have already been widely documented and used.

Confusion matrices, also called error matrices, have been a widely used method for comparing different sets of data. Given a set of C classes $\{c_i\}_{i\in[\![1;C]\!]}$ and two different classifications of a same set of objects, the error matrix is noted $A = (a_{i,j})_{i,j\in[\![1;C]\!]}$. Element $a_{i,j}$ represents the number of objects that were assigned to class c_i by the first classifier and to class c_j by the second.

If both classifications are perfectly agreeing, all but the diagonal elements would be 0. From a qualitative point of view, a good classification agreement would correspond to an 'nearly' diagonal matrix. Having comparably high coefficients $a_{i,j\neq i}$ means that there is a comparably high confusion between classes c_i and c_j.

[2] More sophisticated or accurate methods have also been presented, at the cost of a much higher theoretical complexity [Efron, 1984; DiCiccio and Efron, 1991, 1996] and are limited to specific cases.

When validating maps obtained by remote sensing data, error matrices are commonly used to compare classification results with ground truth [Congalton, 1991, 2001; Foody, 2002; Smits et al., 1999; Stehman and Czaplewski, 1998]. The classical 'standard-way' consists of associating lines with classification results while columns code stands for the ground-truth data.

For the comparison of consecutive flight lines, this convention is not applicable anymore: the rows and columns accounts for the classification results for each flight-line. If they are parallel, it is proposed to place the classification of the south-eastern-most flight-line as the matrix-columns and the north-western most as the rows. This allows to preserve potential effects of the acquisition geometry when analysing several sets of overlaps. In case of crossed flight-lines, such a differentiation is not feasible anymore: the representation of the error matrix should be discussed on a case by case basis.

Once an error matrix has been established, several coefficients can be retrieved [Stehman, 1997]. As before, the total number of pairs of pixels is noted N:

$$N = \sum_{i=1}^{C} \sum_{j=1}^{C} a_{i,j}$$

The overall proportion of agreement, also called total accuracy, $p_o = \sum_{i=1}^{C} a_{i,i}/N$, is likely the simplest and most widely used quality assessment inferred from an error matrix.

Congalton [1991] as well as subsequent authors Foody et al. [2013]; Shao and Wu [2008]; Stehman [1997] make use of the so-called user's and producer's accuracy of class c_i: respectively $p_u(c_i)$ and $p_p(c_i)$ (see equation 6.2.11). When comparing an imagery-based classification with ground-truthing data, the user's accuracy is the probability of a pixel classified as c_i does indeed correspond to a class c_i in the ground-truthing data. On the other hand, the producer's accuracy accounts for the probability of a ground-truth pixel of class c_i to be correctly classified [Congalton, 2001].

$$\begin{aligned} p_u(c_i) &= a_{i,i} / \sum_{j=1}^{C} a_{i,j} \\ p_p(c_i) &= a_{i,i} / \sum_{j=1}^{C} a_{j,i} \end{aligned} \tag{6.2.11}$$

Although the distinction between producer's and user's accuracy makes sense when comparing an imagery-based classification with ground-truthing data, this is less understandable when dealing with a cross-comparison of two flight-lines as performed here: there is no a priori distinction between each flight-line. On the other hand, asymmetries in the cross-confusion matrix indicate classification biases and should be explicitly described. It is therefore proposed to make use of the relative symmetric accuracy $p_s(c_i)$ and the relative asymmetric accuracy, $p_a(c_i)$ as defined in equation 6.2.12.

$$\begin{aligned} p_s(c_i) &= p_u(c_i) + p_p(c_i) \\ p_a(c_i) &= p_u(c_i) - p_p(c_i) \end{aligned} \tag{6.2.12}$$

The relative symmetric accuracy accounts for the overall 'average' agreement of both flight-lines regarding class c_i. If classification was achieved in perfect agreement, it would be 2. On the other hand, in case of 'perfect' disagreement, it would be 0. The asymmetric accuracy accounts for the discrepancies between both flight lines. If there is no asymmetry, it should be zero, while values close to one (resp. minus one) indicate that there are strong disagreement between the eastward and the westward flight-line classifications.

It has been recommended that, besides a classical error matrix analysis, an error matrix standardisation should take place [Congalton, 1991]. Such a matrix scaling - also called matrix ranking - basically scales the error matrix so that the sums of its rows (rows marginals) and the sums of its columns (columns marginals) are all equal to one. This is typically done by applying the iterative proportional fitting procedure (IPF) [Pukelsheim and Simeone, 2009]. Stehman [2004] has however presented several empirical and theoretical arguments against and strongly advocates against its use in remote sensing. One of the main reasons consists of the choice of the marginals. Setting all marginals to one assumes a uniform homogeneous distribution of all classes which is, in practice, almost never the case. Adjusting the marginals to an actual a priori distribution requires on the other hand an exhaustive prior knowledge of the mapped area. Such a knowledge would make the interest of a mapping task comparably null and void except for a very few specific cases.

Continuous Variables

For data types accepting computations of mean and variances, hypothesis testing methodologies can be used to test whether mapping results are agreeing between both flight-lines. This is typically done by means of a paired t-test on the mapping results means [NIST, 2013].

The null hypothesis is the equality of the mapping means obtained for each flight-line. Writing $(m_k, \tilde{m}_k)$ the mapping associated with pairs of observations, let the empirical mean and standard deviation of the discrepancies be defined as follows:

$$\begin{cases} \bar{m} &= \frac{1}{N}\sum\limits_{k=1}^{N} m_k - \tilde{m}_k \\ \bar{\sigma} &= \sqrt{\frac{1}{N-1}\sum\limits_{k=1}^{N}\left(\bar{m} - m_k + \tilde{m}_k\right)^2} \end{cases} \tag{6.2.13}$$

The test statistics can then be defined as being $t = \sqrt{N}\bar{m}/\bar{\sigma}$. The degree of freedom is then $\nu = N - 1$. Once a significance level, α or a confidence coefficient $1 - \alpha$ and the associated $100\,(1-\alpha/2)^{th}$ percentile of the t-distribution - $Q_{1-\alpha/2}\left(\mathcal{T}_\nu\right)$ - have been computed, the null hypothesis will be rejected if [NIST, 2013]:

$$|t| \geqslant Q_{1-\alpha/2}\left(\mathcal{T}_\nu\right) \tag{6.2.14}$$

The test's actual result should be stressed [Nuzzo, 2014]. A rejection of the null-hypothesis - the mean equality - with a significance level α does state that there is a risk of α that the means are actually equal. On the other hand, accepting the null-hypothesis, 'only' states that there is no evidence against it. In other words, rejecting it means that the probability of observing a given data-set, given the hypothesis, is sufficiently low to be neglected. It does in particular not assess the probability of the null-hypothesis given the data [Falk and Greenbaum, 1995]. Furthermore, a tiny difference of the means might lead to a rejection, although it is not significant in an engineering sense.

These limitations and the overrating of the hypothesis testing methodologies are frequently stressed [Anderson et al., 2000; Falk and Greenbaum, 1995; Gigerenzer, 2004; Nuzzo, 2014] although hypothesis testing remains a valuable tool [Morey et al., 2014].

Nuzzo [2014], in particular, recommends providing the confidence interval which corresponds to the recommendations from QA4EO. For nominal data however, confidence intervals make little sense, all the same for ordinal data. In such cases, one has to rely solely on AD and MD, along with the bootstrap of their sampled distributions.

Class	p_u	p_p	p_a	p_s	**Class**	p_u	p_p	p_a	p_s
2	0.727	0.414	0.313	1.141	3	0.547	0.779	−0.232	1.326
4	0.496	0.521	−0.024	1.017	5	0.569	0.451	0.118	1.020
6	0.795	0.751	0.044	1.547	7	0.576	0.661	−0.085	1.238
8	0.611	0.709	−0.098	1.320	9	0.504	0.738	−0.233	1.242
10	0.856	0.906	−0.049	1.762	11	0.109	0.250	−0.141	0.359
12	0.608	0.679	−0.071	1.287	13	0.639	0.609	0.030	1.247
14	0.245	0.122	0.123	0.366	15	0.587	0.701	−0.114	1.288
16	0.364	0.173	0.191	0.536	17	0.435	0.196	0.239	0.631
19	0.794	0.560	0.234	1.354	20	0.762	0.772	−0.010	1.534
21	0.356	0.604	−0.249	0.960	22	0.312	0.033	0.280	0.345
23	0.427	0.331	0.097	0.758	24	0.145	0.206	−0.061	0.350
25	0.510	0.594	−0.084	1.104	29	0.250	0.333	−0.083	0.583
30	0.000	0.000	0.000	0.000	31	0.257	0.541	−0.284	0.798
32	0.214	0.188	0.027	0.402	34	0.562	0.617	−0.055	1.179
35	0.451	0.266	0.185	0.717	36	0.170	0.193	−0.023	0.363
38	0.000	0.000	0.000	0.000	41	0.107	0.273	−0.166	0.380
42	0.303	0.186	0.118	0.489	43	0.148	0.121	0.027	0.269
44	0.000	0.000	0.000	0.000	45	0.350	0.370	−0.020	0.719
47	0.387	0.857	−0.470	1.244	48	0.000	N/A	N/A	N/A
49	N/A	0.000	N/A	N/A	51	0.295	0.358	−0.063	0.653
52	0.042	0.167	−0.125	0.208	53	0.204	0.115	0.089	0.319
54	0.588	0.526	0.062	1.115	55	0.400	0.667	−0.267	1.067
56	0.430	0.352	0.078	0.782	57	0.397	0.346	0.051	0.743
59	0.460	0.504	−0.044	0.964	60	0.286	0.133	0.152	0.419

Table 6.3: Symmetric and asymmetric accuracies: p_s and p_a along with so-called producer's and user's accuracies p_p and p_u for classification results between all flight-lines: Sokolov 2010 data-set. Classes with absolutely no matches were removed from the table.

6.2.4 Case Study: Nominal Data

This section reviews illustrative results based on nominal data, namely the classification of the Sokolov 2010 data-set as described in section 6.1.2.

Raw Results

In order to facilitate an interpretation, and due to the large amount of classes, a graphical representation of the global relative error matrix for the Sokolov 2010 classification is provided on figure 6.3. It makes uses of a logarithmic scale, where a number N of occurrences, is represented by $\log(N+1)$. Detailed relative cross-confusion matrices are provided in appendix B in tables B.1 to B.6. The total matrix over the six overlaps is provided in tables B.7 and B.8.

The detailed symmetric and anti-symmetric proportions are listed in tables B.9 and B.14. Aggregated results for all flight-lines are listed in table 6.3.

Analysis

Several comments have to be made. To start with, some classes are not present in the overlapping areas. This is for example the case of the 'vegetated water' (index 18) which can therefore not be analysed. This is one of the main drawbacks of the cross-comparison methodology.

The overall proportion of agreement between two flight-lines is 69%. The mapping was however designed to discriminate between different minerals and little attention was paid for vegetation discrimination, hence a higher failure rate for these classes. When discarding vegetation classes from the computations, the overall agreement increases up to 86%, corresponding to the classical 85% accuracy requirement.

A closer look at figure 6.3 reveals several other features. Most of the greatest mis-classifications are actually occurring between different 'generic' (or back-ground) land-covers. For example classes 15 (soils) or 20 (vegetation type 2) correspond to a large proportion of pixels: each of them corresponds to more than 10^5 pairs of compliant pixels, thus covering more than 1.6 km^2 of the overlapping parts of the flight-lines. They can therefore be seen as background - 'by default' - land-covers.

Since they correspond to the background classes, they are also more prone to include 'borderline' pixels containing a mixture of land-covers. On the other hand, pairs of pixels, although 'mostly' corresponding one to another, are unlikely to cover perfectly overlapping areas. These footprint discrepancies in turn leads to different percentages of land-cover types within each pair of border pixels, hence potentially leading to different classifications. This explains the high mismatches observed for these generic land-classes.

Despite being a quite straightforward explanation for many mis-classifications - e.g. 'coal type 1' (index 3) with 'soil' (index 15), the previous analysis fails to account for observed asymmetries which especially occurs on some vegetation classes. For example class 'vegetation type 1' (index 19) and 'vegetation type 3' (index 21) both have asymmetries above 0.23 in absolute value. This corresponds in both cases to a difference of factor two between the so-called producer's and user's accuracies. This asymmetry can be explained by the fact that vegetation is very prone to exhibiting strong BRDF effects, thus making its spectral analysis in presence of consequent changes of viewpoints difficult [Asner et al., 1998; Bousquet et al., 2005; Comar et al., 2012; Dorigo, 2008; Peltoniemi et al., 2005; Solheim et al., 2000].

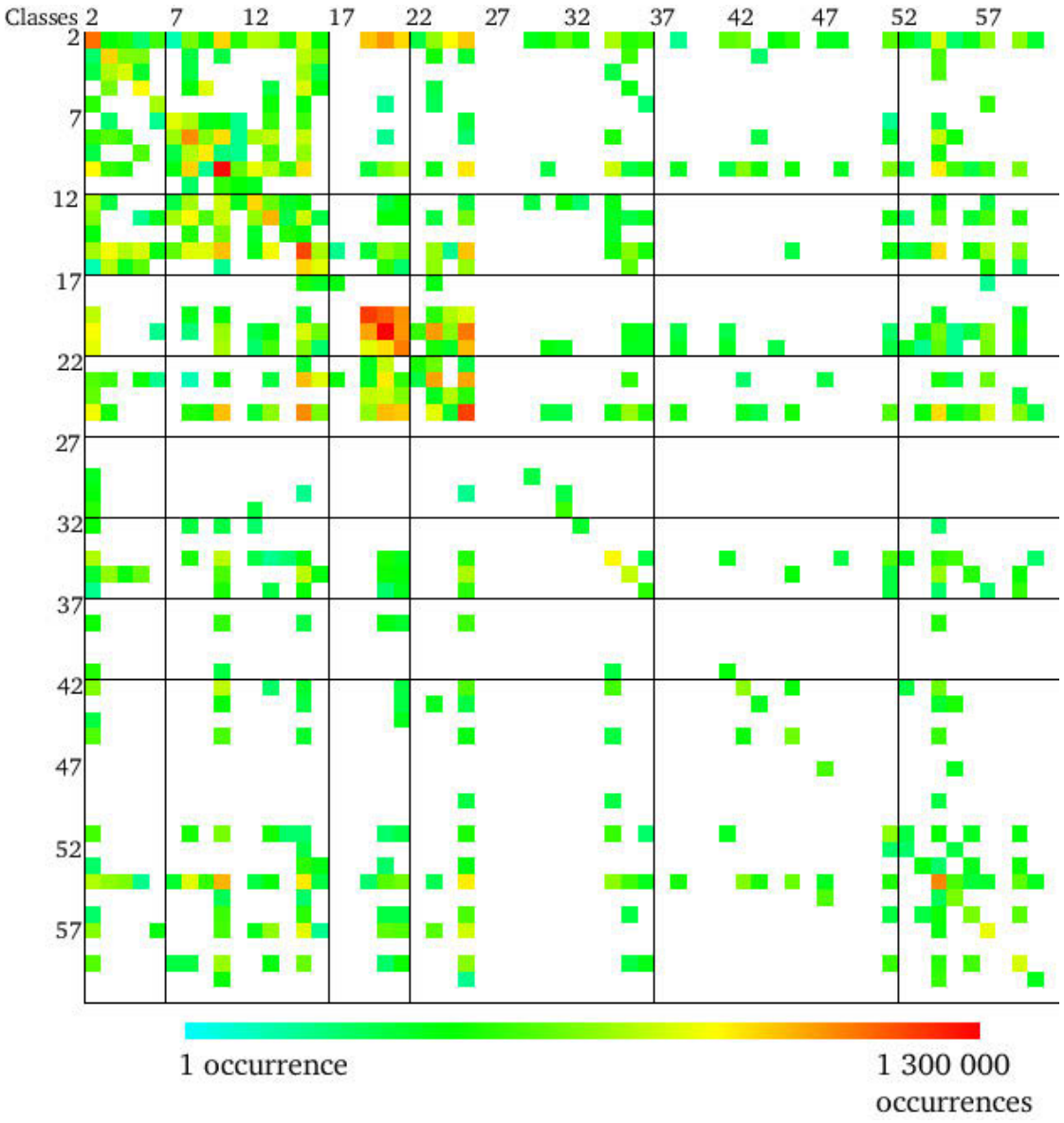

Figure 6.3: Relative confusion matrix: end-members from table 6.2). Columns: eastern flight-lines, Rows: western flight-lines - logarithmic scale, white: no occurrence.

Opening

A preliminary summary can start by examining results obtained for AD as displayed on table 6.4. Summarised results are quite close to 0, indicating a good reliability of the classification. The 'vegetation-BRDF-issue' is furthermore also indicated by AD: the best score is indeed achieved between flight-lines 2 and 3, which are also the ones exhibiting the more open soils and areas with fewer vegetation.

Flight-Lines	AD	$\overline{AD}$	99% CI
# 1 vs. # 2	$1.768\ 10^{-3}$	$1.768\ 10^{-3}$	$[1.759; 1.777]\ 10^{-3}$
# 2 vs. # 3	$1.638\ 10^{-3}$	$1.638\ 10^{-3}$	$[1.624; 1.652]\ 10^{-3}$
# 3 vs. # 4	$2.395\ 10^{-3}$	$2.395\ 10^{-3}$	$[2.373; 2.416]\ 10^{-3}$
# 4 vs. # 5	$4.649\ 10^{-3}$	$4.649\ 10^{-3}$	$[4.578; 4.721]\ 10^{-3}$
# 5 vs. # 6	$9.442\ 10^{-3}$	$9.440\ 10^{-3}$	$[9.102; 9.769]\ 10^{-3}$
# 6 vs. # 7	$2.242\ 10^{-2}$	$2.241\ 10^{-2}$	$[1.968; 2.496]\ 10^{-2}$

Table 6.4: Example of AD values (raw and bootstrap) along with confidence intervals for the Sokolov 2010 data-set (classification).

One can therefore notice that the produced results are concordant with the expected qualitative variability of the observed data. It should however be noted that, for nominal data, MD would only take the value 1 - provided there is at least one single classification discrepancy. It is - in this case - therefore not really worth investigating.

The presented results however require a lot of investigation: values obtained for AD as well as p_a and p_s are only usable when further manual analysis is performed.

6.2.5 Case Study: Absolute Data

As a second example, one analyses the NDVI distribution between flight-lines 1 to 4 of the Witbank 2013 data-set. Since NDVI corresponds to absolute data, the means of a basic analysis are more restricted. Section 6.2.1 has already shown how the cumulative distribution could be used to retrieve a maximal admissible error, the following therefore focuses on the remaining parameters.

To start with, a survey of the retrieved AD and MD coefficients are displayed on table 6.5. It can be noticed that the retrieved MD is always equal or greater to its bootstrap confidence interval: this is a logical behaviour since it is taken as being the maximum discrepancy. All bootstrap re-samples are therefore smaller.

Flight-Lines	AD	$\overline{AD}$	AD: 99% CI	MD	$\overline{MD}$	MD: 99% CI
# 1 vs. # 2	$5.31\ 10^{-5}$	$5.31\ 10^{-5}$	$[5.27; 5.36]\ 10^{-5}$	0.194	0.0.190	$[0.172; 0.194]$
# 2 vs. # 3	$6.31\ 10^{-5}$	$6.31\ 10^{-5}$	$[6.23; 6.38]\ 10^{-5}$	0.340	0.332	$[0.300; 0.340]$
# 3 vs. # 4	$5.92\ 10^{-5}$	$5.92\ 10^{-5}$	$[5.85; 5.98]\ 10^{-5}$	0.532	0.523	$[0.482; 0.532]$

Table 6.5: Example of AD and MD values (raw and bootstrap) along with confidence intervals for NDVI over the Witbank 2013 data-set.

As a further test, the mean NDVI discrepancy between each flight-line was tested for equality: the corresponding t values and quantiles are provided on table 6.6. As one can see, the hypothesis of an NDVI equality for both flight-lines can be rejected. This is qualitatively visible when looking at the distribution DD (see figure 6.4) and observing that the distributions are indeed not centered on zero, hence a manual validation of the automated test results.

Flight-Lines	t	$\mathcal{Q}_{0.995}\left(\mathcal{T}_{N-1}\right)$	**Result**
# 1 vs. # 2	250.5	2.576	Failed
# 2 vs. # 3	176.9	2.576	Failed
# 3 vs. # 4	174.0	2.576	Failed

Table 6.6: Example of testing for NDVI values equalities: Witbank 2013 data-set.

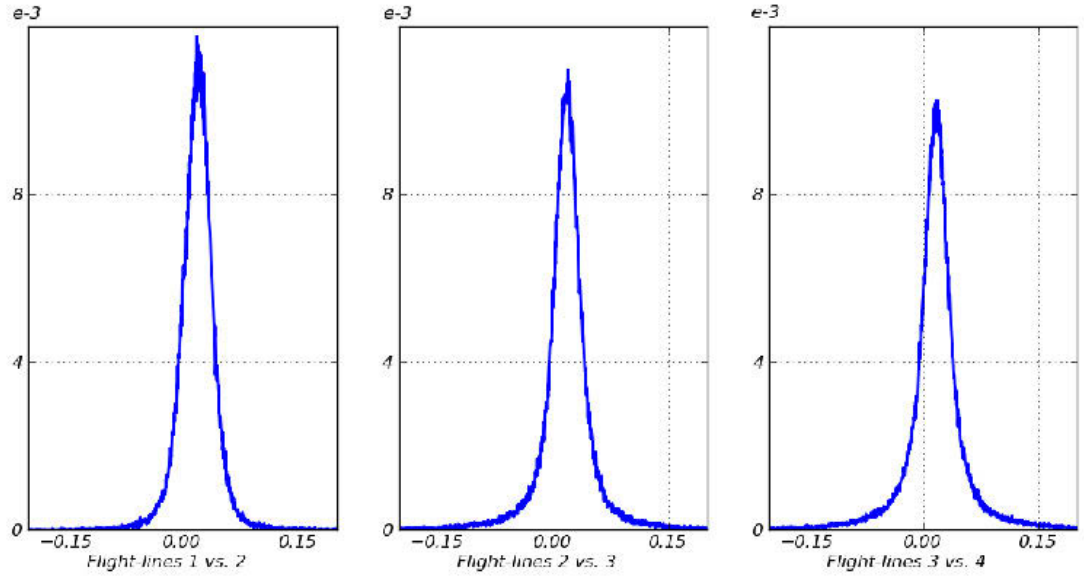

Figure 6.4: Example of DD plots for the NDVI computations: Witbank 2013 data-set.

6.2.6 Conclusion

Although it requires extensive interpretation, the basic indicators presented in this section allow a 'data-near' investigation of the results. These, except for AD, are however dependent on the data type used within the mapping application. The latter, along with the similar MD and CD, is therefore preferred as being more generic. Section 6.3 presents some more standardised upper-level quality indicators.

These indicators can - and should - finally be linked with other potential quality indicators. When ground truth data is available, even in limited number, it can be used to estimate the bias of the L3 product. On the other hand AD, MD and CD account for the errors present in the product, hence complementing the latter assessment. They can however be computed on virtually any land-product, while ground-truth might be unavailable.

6.3 In-depth Analysis: Cohen's Kappa Coefficient

The previous sections provided several tools to analyse the results based on cross comparisons of mappings from both flight-lines. These are however quite crude in the sense that they require further analysis before being of any use. This section presents some higher-level quality indicators, allowing for a more direct evaluation of their results. These are called reliability or inter-rater agreement coefficients. Strictly speaking and following Krippendorff [2004] definitions, the agreement is what is observed, while reliability is what is inferred from these observations. This semantic nuance has however little importance for this work: In the following sections, both words being used in an interchangeable fashion, the actual discrepancy between both notions being explicitely addressed.

To start with, section 6.3.1 reviews the main properties of the well known and almost standard Cohen's κ coefficient. Section 6.3.2 in turn reviews other coefficients, some of which were seemingly neglected in the fields of geosciences. At last, section 6.3.3 presents the corresponding limitations.

6.3.1 Inter-Rater Reliability: the Kappa Coefficient

A classical alternative to the 'raw' error matrix analysis consists in the computation of the so-called κ coefficient. Its use has often been been considered as a must to assess the quality of a mapping product [Congalton, 1991, 2001; Foody, 2002, 2010; Foody et al., 2013; Lu et al., 2008; Næsset, 1996; Smits et al., 1999; Stehman, 1996; Schiewe and Kinkeldey, 2013; Wilkinson, 2005] and others.

In order to understand the underlying principle, imagine a 2-class classification scheme, that would randomly associate a class to each pixel with the probability 99% for the first one, and 1% for the second. The overall agreement of two consecutive runs would then be $0.99 * 0.99 + 0.01 * 0.01 \approx 0.98$. A naive interpretation would conclude that this is the proof of the method's reliability. In practice, however, this result has nothing to do with the actual classification reliability. Cohen's kappa coefficient, noted κ, aims at tackling this issue by accounting for the 'by chance' agreement.

The κ coefficient was first formalised by Cohen [1960] in the field of medical diagnostics, the goal being to compare diagnostics made by two physicians. It has since been commonly used for validating remote sensing data [Congalton, 1991, 2001; Næsset, 1996; Stehman, 1996].

The main idea consists of analysing the proportion of mismatches, just as for an error matrices, but taking into account the number of agreements occurring by chance. Given two methods of evaluation, and noting p_o the proportion of measures where both are agreeing, and p_c the proportion where the agreement occurs by chance, the kappa coefficient is defined as equation 6.3.1.

$$\kappa = \frac{p_o - p_c}{1 - p_c} \tag{6.3.1}$$

It can be directly derived from equation 6.3.1 that if both methods strictly agree, then $\kappa = 1$. On the other hand, if agreement is lower than the by chance agreement, κ will be negative. Cohen [1960] considers the probability of by chance agreement as being the joint probability of the marginal proportions for each class. Given C classes and N elements to be classified, the corresponding error matrix is written, $A = (a_{i,j})_{i,j \in [\![1,C]\!]}$. The p_o and p_c are therefore defined as in equation 6.3.2.

$$\begin{cases} p_o = \frac{1}{N} \sum_{k=1}^{C} a_{k,k} \\ p_c = \frac{1}{N^2} \sum_{k=1}^{C} \left(\left(\sum_{l=1}^{C} a_{l,k} \right) \cdot \left(\sum_{l=1}^{C} a_{k,l} \right) \right) \end{cases} \tag{6.3.2}$$

Its variance can be estimated using equation 6.3.3 [Rossiter, 2004; Fleiss et al., 2003] or, for $N \gg 100$ the corresponding approximation [Cohen, 1960].

$$\begin{aligned} \tilde{\sigma}_\kappa^2 &= \frac{1}{N}\left(\frac{p_o(1-p_o)}{(1-p_c)^2} + \frac{2(1-p_o)(2p_o p_c - \phi_1)}{(1-p_c)^3} + \frac{(1-p_o)^2\left(\phi_2 - 4p_e^2\right)}{(1-p_c)^4}\right) \\ &\approx \frac{1}{N}\left(\frac{p_o(1-p_o)}{(1-p_c)^2}\right) \\ \text{with} & \\ \phi_1 &= \frac{1}{N^2}\sum_{k=1}^{C} a_{k,k}\left(\sum_{l=1}^{C} a_{k,l} + a_{l,k}\right) \\ \phi_2 &= \frac{1}{N^3}\sum_{k=1}^{N}\sum_{l=1}^{C} a_{k,l}\left(\sum_{m=1}^{C} a_{m,k} + a_{l,m}\right)^2 \end{aligned} \tag{6.3.3}$$

Under the Gaussian assumption, the following 'classical' confidence limits can be retrieved from equation 6.3.3:

- 95% confidence limits: $\kappa \pm 1.96\tilde{\sigma}_\kappa$
- 99% confidence limits: $\kappa \pm 2.58\tilde{\sigma}_\kappa$

Cohen [1960] also proposes to test for significance by referring the critical ratio $\kappa/\tilde{\sigma}_\kappa$ to the normal curve, the null hypothesis to hopefully be rejected being $\kappa = 0$.

Finally, Cohen [1960] provides an estimation of the '*maximum*' κ: κ_m. The term maximum is here to be understood as maximum given the marginals. In other words, $1-\kappa_m$ is the proportion of agreement that cannot be achieved due to differing marginals when chance agreements are excluded.

κ_m can therefore be obtained by replacing, in equation 6.3.1, p_o by p_m:

$$p_m = \frac{1}{N}\sum_{k=1}^{C} \min\left(\sum_{l=1}^{C} a_{k,l}, \sum_{l=1}^{C} a_{l,k}\right) \tag{6.3.4}$$

Several drawbacks are however associated with the κ values. First of all, values taken by κ are not scaled: given a κ value, it is not possible to state whether it is a good or bad agreement. Although Landis and Koch [1977] have proposed such an interpretation scale, most authors agree that it is not valid in the generic case [Artstein and Poesio, 2008; Gwet, 2002; Hsu and Field, 2003; Sim and Wright, 2005; Uebersax, 1987; Viera and Garrett, 2005; Wongpakaran et al., 2013]. The main argument is based on the fact that κ is actually affected by prevalence of the observations, hence preventing its comparison between two consecutive studies. Consider for example the two fictional cross confusion matrices presented in table 6.7: in terms of reproducibility, a higher score would have been expected in the first case while the opposite actually occurs. These reasons leads to reconsider the adequacy of the use of κ as a standard coefficient for comparing classification results. Section 6.3.2 presents some alternatives.

Case 1			Case 2		
	A	B		A	B
A	18	8	A	18	16
B	8	20	B	0	20
$\kappa \approx 0.41$			$\kappa \approx 0.46$		

Table 6.7: Example of a 'failure' of κ caused by asymmetry.

6.3.2 Other Coefficients

Although κ is the reference inter-rater agreement metric for geographical applications, several alternatives or enhancements have been proposed.

An Extension of κ: the Weighted κ

Although not as commonly spread as the classical κ coefficient, the weighted kappa, or κ_w coefficient enables to account for the relative importance of discrepancies [Cohen, 1968; Fleiss et al., 2003]. This allows, for example, to consider that a mis-classification between two vegetation classes is less problematic than a mis-classification between vegetation and urban classes. This can be achieved by associating a weight matrix $W = (w_{k,l})$ to the error matrix A such as for all (k, l) from $[\![1; C]\!]$:

- $w_{k,l} \in [0; 1]$
- $w_{k,k} = 1$
- if $k \neq l$, then $w_{k,l} < 1$

If such properties are met, $p_o(\kappa_w)$ and $p_c(\kappa_w)$ can be defined as in equation 6.3.5. When W is the identity matrix, this yields the usual κ back. If both raters are considered as equivalent, W has to be symmetric. On the other hand, asymmetry between raters could be modelled by introducing asymmetries in W.

$$\begin{cases} p_o(\kappa_w) &= \frac{1}{N} \sum_{k=1}^{C} \sum_{l=1}^{C} w_{k,l} a_{k,l} \\ \\ p_c(\kappa_w) &= \frac{1}{N^2} \sum_{k=1}^{C} \sum_{l=1}^{C} w_{k,l} \left(\sum_{m=1}^{C} a_{k,m} \right) \left(\sum_{m=1}^{C} a_{m,l} \right) \end{cases} \tag{6.3.5}$$

Just as for κ, Cohen [1968] and Fleiss et al. [2003] provide an estimator of the variance of κ_w for large sample sizes:

$$\tilde{\sigma}^2_{\kappa_w} \approx \frac{1}{N\left(1 - p_c\right)^2} \sum_{k=1}^{C} \sum_{l=1}^{C} \left[\sum_{m=1}^{C} a_{m,l} \right] \left[\sum_{m=1}^{C} a_{k,m} \right] \left[\frac{w_{k,l} - \hat{w}_k - \check{w}_l}{N} \right]^2 - p_c^2 \tag{6.3.6}$$

Where p_c is the chance rate as defined in equation 6.3.5 and

$$\hat{w}_k = \sum_{l=1}^{C} \frac{w_{k,l}}{N} \sum_{m=0}^{C} a_{m,l} \quad \text{and} \quad \check{w}_k = \sum_{l=1}^{C} \frac{w_{l,k}}{N} \sum_{m=0}^{C} a_{l,m}$$

As for the classical κ, the critical ratio $\kappa_w / \tilde{\sigma}_{\kappa_w}$ can be used to test results for significance.

Fleiss and Cohen [1973] showed that κ_w is actually a generalisation of the well-known intra-class correlation coefficient. The latter does indeed correspond to the case where weights are given by

$$w_{k,l} = 1 - \left(\frac{k - l}{C - 1} \right)^2$$

Despite being considered as more suitable than the simple κ coefficient [Artstein and Poesio, 2008], the weighted κ requires setting up the weight matrix, which requires some in-depth information not necessarily available.

Other Kappa-like Coefficients

Definitions other than the one from equation 6.3.2 can of course be chosen for computing the by chance agreement proportion p_c. That could in turn be used within equation 6.3.1. To start with, the Scott's π coefficient is strictly identical to the κ coefficient, except for the computations of p_c. Scott's π does assume that the distribution of classifications by both methods is the same and therefore estimates p_c using the sum of marginal proportions: see equation 6.3.7 [Scott, 1955].

$$p_c(\pi) = \frac{\sum_{k=1}^{C} \left(\sum_{l=1}^{C} a_{k,l} + a_{l,k} \right)^2}{4N^2} \tag{6.3.7}$$

Scott's π has actually been extended to the use case where more than 2 observers are compared, leading to the so-called Fleiss K [Fleiss, 1971]. Given M observers, or raters, and noting $(c_{k,l})_{k \in [\![1,N]\!], l \in [\![1,C]\!]}$ the number of raters that classified the k^{th} element as being from class l, $p_o(K)$ and $p_c(K)$ are defined as in equation 6.3.8.

$$\begin{cases} p_o(K) &= \frac{1}{NM(M-1)} \sum_{k=1}^{N} \sum_{l=1}^{C} \left(c_{k,l}^2 - NM \right) \\ \\ p_c(K) &= \frac{1}{NM} \sum_{k=1}^{C} \left(\sum_{l=1}^{N} c_{l,k} \right)^2 \end{cases} \tag{6.3.8}$$

This family aside, the $\mathcal{S}$ coefficients initially introduced by Bennett et al. [1954] and reinvented by several other authors [Artstein and Poesio, 2008; Hsu and Field, 2003] consider that chance agreement is uniformly distributed among all raters and all classes. This yields $p_c = 1/C$.

In remote sensing, another commonly used choice for the p_c term consists of introducing a priori probabilities to the category memberships: p_k for class k. The κ coefficient is then renamed as τ-index [Ma and Redmond, 1995; Næsset, 1996]. In practice, this knowledge is however usually not available: the prior probabilities may then be assumed to be uniformly distributed: $\forall k$, $p_k = 1/N$ [Rossiter, 2004]. In the latter case, τ becomes an asymmetric mixture between π and $\mathcal{S}$ coefficients.

$$p_c(\tau) = \sum_{k=0}^{C} \left(p_k \sum_{l=0}^{C} a_{l,k} \right) / N \tag{6.3.9}$$

Both $\mathcal{S}$ and π are making stronger assumptions on the distributions and the relationships of by chance classifications than the κ coefficient. For example, Scott's π assumes that both by chance classifications are made using the same underlying probability distribution. When comparing a ground-based classification and an image-based one, this might seem a quite constraining hypothesis, thus leading to a preferred use of κ. On the other hand, when comparing classifications performed over two consecutive flight-lines, this hypothesis makes more sense.

Beyond Nominal Data

All the coefficients presented so far are only dealing with nominal data. Several other types of coefficients have however been developed for different data types. Kendall rank coefficient, also called 'Kendall's τ can be used on ordinal data: it compares the correlation between the orderings of the data. In a similar manner, Kendall's W accounts in a slightly more refined manner for the case where more than two observations have to be compared [Kendall and Smith, 1939]. The Goodman and Kruskal λ for nominal data and their γ for ordinal data take into account probabilistic models for activity [Goodman and Kruskal, 1954, 1959].The Pearson product moment correlation coefficient, also called Pearson correlation - or ρ - is likely the best known two-rater agreement. It accounts for a linear relationship when both variables are normally distributed. Pearson's correlation can be computed from interval data up to absolute data.

The Spearman's r, or Spearman's rank correlation coefficient can be used to assess dependence between two populations in terms of a monotonic function. It can be used from ordinal data up to absolute data. It actually consists of a translation of pairs of values into ranks before computing the Pearson's correlation coefficient between these. The Cronbach's α has been designed for ratio data and quantifies the consistency by which observation units are judged by observers. It however does not account for how many observations actually agree [Hayes and Krippendorff, 2007].

6.3.3 Limitations

However, among all these coefficients - including the Cohen's κ, none is able to cover indistinctly all data types: they are therefore lacking generality. Furthermore, several of them, including the κ coefficient, are not able to tackle cases where data is rated by more than two observers or when some ratings are missing. Although the later case could potentially be tacked by adding a 'no data' class or category, this would however be likely to introduce some bias in the computations and should therefore be avoided.

In practice both cases are likely to happen when working on Earth observation data. For example, when dealing with two overlapping flight-lines, it might happen that some 'ground-truth' data is actually available within the overlap areas. In such cases, it would be annoying to discard it from computations, just because one cannot deal with it.

One can consider the case where the mapping application includes a 'cannot-judge' decision step. This is for example the case of the classification scheme presented in section 6.1.2. A potential solution to both issues might reside in the so-called Krippendorff's α coefficient [Artstein and Poesio, 2008].

6.4 A Unified Framework: Krippendorff's α

This section introduces a more generic inter-rater reliability coefficient, Krippendorff's α, which overcomes many limitations of the aforementioned coefficients while being generic a the same time [Krippendorff, 1970, 2004; Hayes and Krippendorff, 2007]. Although its use has been recommended in various fields such as social sciences [Krippendorff, 2004], medicine [Wongpakaran et al., 2013] or linguistics [Artstein and Poesio, 2008], it does not seem to have ever been introduced in geosciences where the use of κ and κ-like coefficient does not appear to have been questioned except on few occasions [Olofsson et al., 2013] where no alternative was proposed.

Section 6.4.1 reviews some of the properties one could expect from a reliability coefficient. Sections 6.4.2 to 6.4.4 then presents the main computational aspects before making the link with other coefficients in section 6.4.5.

6.4.1 Motivations

Instead of producing yet another reliability coefficient, this section starts by reviewing what a reliability coefficient should account as explicated in [Hayes and Krippendorff, 2007; Krippendorff, 2004, 2011b]:

1. It should be able to account for the amount of agreement between at least two observers. This also implies that the amount of observers and their permutation should not affect the data. One could note that this is in slight disagreement with more classical frameworks applied to Earth-observation, where observers are not interchangeable (τ-index, etc). Observers are supposed to be independent.
2. It should be based solely on the actual observation, not on the expected categories.
3. It should be scaled with at least two points of defined reliability.
4. It should be able to handle different data types[3] and be modified only by changes to the subsequent metrics.
5. It should have a computable or, ideally, a well known sampling behaviour.

Bearing this five requirements in mind, the principles of Krippendorff's α coefficient can be explained. One should also note that the previously stated coefficients are violating one or several of these conditions, if only be because they can usually only apply to specific cases. Moreover, remote sensing coefficients derived from κ are typically un-compliant with condition one.

6.4.2 Principles

Krippendorff's α general expression is given in equation 6.4.1, where D_o is the observed disagreement among values, and D_e is the disagreement expected by chance instead of by intrinsic properties of the data [Krippendorff, 2011a]. The strength of the approach, however lies within the computation of the two last terms:

$$\alpha = 1 - \frac{D_o}{D_e} \tag{6.4.1}$$

The computations of D_o and D_e are presented here in the generic case. Specific cases where α can be computed in a simpler manner are explained in Krippendorff [2011a].

[3] In Hayes and Krippendorff [2007] vocabulary: these are called level of measurement.

As before, the amount of observation units is noted N. Observation of unit n leads to J ratings - $\{r_{n,j}\}_j$ - corresponding to the judgement issued by the j^{th} observer. It should be noted that $r_{n,j}$ can potentially take the value '*no judgement issued*'. The number of differing observation values different from the latter - potentially JN - is noted K, $\{v_k\}_k$ being the set of obtained values, still excepting the '*no judgement issued*' one.

To start with, the K by N 'units-by-values' matrix, $M = (m_{k,u})$, with $m_{k,n}$ being the number of judgements equals to v_k obtained for the n^{th} observation:

$$m_{k,n} = Card\left(\{r_{n,j} \mid r_{n,j} = v_k\}_{j \in [\![1;J]\!]}\right)$$

Finally, using the notations:

- $c_n = \sum_{k=1}^{K} m_{k,n}$, the number of judgements actually issued for the observation unit n.
- $d_k = \sum_{n|c_n>1} m_{k,n}$, the number of values v_k, omitting observation units with one or no judgement values.
- $C = \sum_{n|c_n>1} c_n$, the total number of values, omitting observation units with one or no judgement values.

D_o and D_e can finally be defined as follows, μ being the metric from section 6.1.1 adapted to the current type of data:

$$\begin{cases} D_o &= \sum_{n=1}^{N} \frac{1}{c_n - 1} \sum_{k=1}^{K} \sum_{l=k+1}^{K} m_{k,n} m_{l,n} \mu\left(v_k, v_l\right)^2 \\ D_e &= \frac{1}{C-1} \sum_{k=1}^{K} \sum_{l=k+1}^{K} d_k d_l \mu\left(v_k, v_l\right)^2 \end{cases} \tag{6.4.2}$$

It can be observed that the first four requirements stated in section 6.4.1 are met in a straightforward manner. In particular, α is scaled from 0 - the categories are unrelated to the described units - to 1 for a perfect reliability. The coefficient can also be negative, which corresponds to systematic disagreements.

6.4.3 Confidence Intervals

Beside its computational complexity, Krippendorff's α suffers from another drawback: there is no general expression for its confidence interval, thus making it necessary to rely on bootstrapping [Krippendorff, 2006]. This, as seen in section 6.2.2, is not intrinsically complex but would require a large amount of re-computations of the α coefficient: typically $B = 20000$ or more. This would in turn increase the computational cost by as many runs.

On the other hand, relying on bootstrapping discards the need for any assumption on the α's distribution, leading to potentially more accurate information [Hayes and Krippendorff, 2007].

6.4.4 Note on Complexity

Once M has been computed, D_o requires in its basic implementation $\mathcal{O}\left(NK^2\right)$ evaluations of the metric μ, while D_e can be assessed in only $\mathcal{O}\left(K^2\right)$ evaluations. Accounting for bootstrapping, this yields a complexity of $\mathcal{O}\left(NBK^2\right)$. Storing M requires on the other hand $\mathcal{O}\left(NK\right)$ space: it can however be computed only once for each bootstrap iteration.

In practice, the main unknown is K. For a classification of ordinal data, it is typically smaller than 100 and can even be smaller than 10. On the other hand, when dealing with continuous data, the amount of encountered values can virtually explode: 1 and 0.9999999999999 are two different values, which leads to very high computational times, K having a polynomial impact. The last issue can however be partially left aside by 'sub-sampling' the results to quantified categories, at the cost of a loss of information.

One can also note that, in comparison to the Cohen's κ, the α coefficient involves slightly more abstract equations as well as a much higher computational complexity.

6.4.5 Relationships Between Krippendorff's Alpha and Other Coefficients

Despite having a more complex form, Krippendorff's α can be simplified in some specific cases. When dealing with nominal data and when all observation units lead to a statement (i.e. no 'no judgement issued' class), Scott's π and Fleiss' K are approximations of Krippendorff's α as long as N is large. The coefficients are actually asymptotically equal when the number of observations tends to infinity.

For ordinal data and in the absence of ties and as long as all observation units lead to a statement, Spearman's ρ is also an approximation of α, provided that N is large: this becomes an equality when N tends to infinity.

Finally one should stress the differences between κ and α. Besides its much wider scope, α can be related to Scott's π, assuming a similarity between both coders. This is in accordance with point one from section 6.4.1 and achieved by analysing the coincidences of observations. On the other hand, Cohen's κ fails to account for the unequal classes uses made by the two judges [Krippendorff, 2004].

6.4.6 Conclusions on Theoretical Background

It should be noted that most of the rater agreements usually consider each set of observations as being equivalent. This is interchangeable: taking A^t instead of A would not alter the results. This is in compliance with the study areas where most of the research on the topic was done. This encompasses medical diagnostics, where diagnostics of two or more physicians are compared (e.g. [Wongpakaran et al., 2013]), poll-surveying and marketing where the contribution of different interviews are compared (e.g. [Feng, 2012]), content analysis where different interpreters transcribe information into nominal categories (e.g. [Artstein and Poesio, 2008]), and many other cases.

In remote sensing, however, this symmetry has been left aside, one observation being usually considered as 'the truth' (usually the 'ground-truth') and the second one being the data to be validated (usually, mapping derived from the remote-sensing data) [Congalton, 2001; Foody, 2002; Smits et al., 1999]. This culminates with the τ-coefficient which is intrinsically asymmetric. In the following, this asymmetry is ignored. Comparison, is indeed made between two consecutive flight-lines, neither of which is intrinsically better than the other. This leads back to the roots of inter-rater evaluation.

Despite being considered as almost a standard, the κ coefficient is arguable [Gwet, 2002]. One of the most direct criticisms regards its inability to assess non-nominal data. This is in practice quite a limitation since classification only represents a subset of the available information extraction methods. For example, band ratios such as the widely used NDVI (see section 2.4.2) as well as quantitative analysis [Bayer et al., 2012; Ehrler, 2014] are producing continuous, real valued results: κ coefficient is not applicable in such use case.

In order to address this issue, the use of the Krippendorff's α coefficient is proposed not only for this relative assessment, but also as a replacement for Cohen's κ. Krippendorff's α does indeed tackle several flaws inherant to the κ coefficient while being at the same time extremely versatile. This does of course involve a greater computational cost - especially due to the requirement of a bootstrap to obtain its confidence interval. In the context of this study, where processing durations are of little importance, this drawback is however of little importance.

Notwithstanding criticisms, results based on Cohen's κ coefficient are also presented, hence making the link with this common quality assessment tools. Their meaning and should however not be over-rated.

6.5 Experimental Results

To start with, a look back at the classification results from section 6.1.2 has to be done: the kappa coefficient was computed for each pair of flight-lines, as well as for the combinations of overlaps (see table 6.8). As one can see, the results are not self-compliant: κ coefficients obtained for each flight-line should actually have been included within the confidence intervals obtained for the global data-set. This issue might actually be linked to a violation of the normality hypothesis issued for estimating κ's confidence interval (see section section 6.3.1)[4].

Another explanation can be retrieved from the data itself. One shall remember that the evaluated classification was primarily designed for mineral mapping, little care being paid to the classification of vegetation types. It follows, as can be already seen in figure 6.3, that the vegetation classes are less accurately mapped than the mineral ones. The corresponding κ should therefore also vary depending on the ground-cover type. Although this trend can be qualitatively observed - κ is indeed higher for flight-lines 2 and 3, which corresponds to the overlap with the more open soils - this seems unlikely to be the only valid explanation. Indeed, flight-lines 5, 6 and 7 are sharing the same land-covers, and should therefore exhibit a similar κ, which is in practice not the case.

Flight-Lines	κ **coefficients**	99% **confidence**
#1 vs. #2	0.59	$\pm 4.4\ 10^{-3}$
#2 vs. #3	0.69	$\pm 4.1\ 10^{-3}$
#3 vs. #4	0.66	$\pm 6.2\ 10^{-3}$
#4 vs. #5	0.55	$\pm 1.4\ 10^{-2}$
#5 vs. #6	0.60	$\pm 2.8\ 10^{-2}$
#6 vs. #7	0.65	$\pm 8.0\ 10^{-2}$
global	0.63	$\pm 5.3\ 10^{-3}$

Table 6.8: κ scores obtained over scene Sokolov 2010 (see section F.2): $w = 10$, $n = 7$.

[4] Another possibility is that each flight-line was out of the 99% interval: this is however quite unlikely: the corresponding probability being $0.01^6 = 10^{-12}$...

In opposition to the κ, the Krippendorff's α coefficient provides much more plausible results (see table 6.9). To start with, although producing quite different results for the flight-lines 5 to 7, the corresponding confidence intervals are widely overlapping, thus indicating similar properties. There are similarly sound overlaps of the confidence interval obtained for the complete scene and the ones obtained for one to one comparison of the flight-lines. A noticeable exception is - as expected - results obtained for flight-lines 2 and 3. Corresponding mostly to open-soils, the classification performed in a much better way, thus leading to a higher α.

It should however be noticed that these outputs indicate quite 'bad' results, most results being lower than 2/3 - the limit for allowing tentative conclusions in social sciences according to [Krippendorff, 2006]. This can be - again - explained by the large proportion of vegetation present within the imagery.

Flight-Lines	α **coefficients**	99% **confidence**
#1 vs. #2	0.62	$[0.60; 0.63]$
#2 vs. #3	0.71	$[0.70; 0.73]$
#3 vs. #4	0.65	$[0.63; 0.67]$
#4 vs. #5	0.58	$[0.55; 0.61]$
#5 vs. #6	0.63	$[0.59; 0.67]$
#6 vs. #7	0.53	$[0.39; 0.66]$
Global	0.63	$[0.62; 0.65]$

Table 6.9: α scores obtained over scene Sokolov 2010 F.2: $w = 10$, $n = 7$.

When computed on mineral classes only (i.e. the first 16 classes), the Krippendorff's α coefficient does in practice lead to much more compliant results. It should however be noted, that the overlap between flight-lines 5, 6 and 7 was discarded, due to the absence of a sufficient number of mineral pixels:

- Between flight-lines 5 and 6, only two mineral classes are present. which is not enough to provide a reliable comparison with other flight-lines overlaps. The α coefficient can however be computed as being equal to 0.95 with a 99% confidence interval of $[0.88; 0.99]$. which indicates a very good agreement for *these two specific classes*.
- Between flight-lines 6 and 7 no mineral class was detected, thus making it impossible to compute the reliability coefficients.

One can finally note that the overlap between flight-lines 4 and 5 was included in the comparisons, although they represent only a subset of the scene's mineral classes corresponding to the overburden soils. In particular, the coal classes are absent. Results are summarised in table 6.10. It can be seen that both coefficients indicate a qualitative increase of reliability. However, κ is still not compliant with itself: the confidence intervals are not overlapping either when comparing flight-line to flight-line, or when comparing with the global score.

On the other hand, Krippendorff's α coefficient provides results compliant with their own confidence intervals. A slight exception should however be noted for the overlap between flight-lines 5 and 6 which presents significantly higher scores. As already presented, this is due to the presence of only a subset of the mineral classes found in other flight lines. In particular most of the similar classes (e.g. coals type 1 and type 2) are not present, thus limiting the risk of mis-classifications and increasing the reliability.

Flight-Lines	α	α**: 99% confidence**	κ	κ**: 99% confidence**
#1 vs. #2	0.74	$[0.71; 0.77]$	0.741	$\pm 8.5\ 10^{-3}$
#2 vs. #3	0.74	$[0.68; 0.79]$	0.758	$\pm 1.1\ 10^{-2}$
#3 vs. #4	0.77	$[0.74; 0.80]$	0.819	$\pm 9.9\ 10^{-3}$
#4 vs. #5	0.78	$[0.73; 0.82]$	0.761	$\pm 4.2\ 10^{-2}$
global	0.78	$[0.75; 0.81]$	0.795	$\pm 5.0\ 10^{-3}$

Table 6.10: α and κ obtained over scene Sokolov 2010 (see section F.2): $w = 10$, $n = 7$. Overlaps between flight-lines 5 to 7 were discarded due the absence of mineral pixels.

6.6 Conclusions

This chapter has shown how pairs of pixels, as retrieved in chapter 4, can be used to gain a generic quality assessment over rasterised land-products. Different quality indicators have been proposed, allowing an evaluation of the mapping results reliability. Among these, the AD and MD coefficients allow a basic but quick evaluation, independently from the data type used. On the other hand, it has been proposed to use Krippendorff's α as a higher-level coefficient in place of the limited but widely-used κ coefficient. Both coefficients, eventually along with the presented refined tools for sub-cases, allow a rating of the mapping reproducibility.

It has also been shown how to combine results from different pairs of flight-lines. This latter ability is however limited by several factors. For an in-depth analysis, each flight-line should present similar land-covers, or at least lead to similar mapping results. Furthermore, while Krippendorff's α is quite stable to scale changes, this is not necessarily the case for simpler coefficients. For example, a classification scheme that has 20 matches with 3 errors for a first overlap and 10 matches with 1 error on a second overlap leads to: $AD_1 = \sqrt{3}/20 = 0.087$, $AD_2 = \sqrt{1}/10 = 0.1$ but for the global overlap of $AD = \sqrt{3+1}/(20+10) = 0.067$ where no direct link to AD_1 or AD_2 could be retrieved. This can be seen as a further argument toward the use of Krippendorff's α.

It should finally be stressed one more time that the proposed assessment is a relative one: in other words, it accounts for errors, not for biases. A systematic disagreement between the 'reality' and the mapping application can indeed not be observed when comparing results from flight-line to flight-line. The corresponding bias can therefore not be accounted for by the proposed methodology. In practice, such a limitation however has a limited impact. Systematic bias can indeed be assessed by a comparison with different measurements such as ground-truth and has already been widely assessed - for example, and among many other approaches in [Olofsson et al., 2013, 2014]. This is therefore left aside in the forthcoming developments.

Nevertheless, the proposed indicators AD, MD or even α can be used as generic extensions to the EUFAR quality indicators. Even if they are more than 'just tags' their results are simple enough to interpret for justifying their use on land-products.

Chapter 7

Improving Land-Products

If you can not measure it, you can not improve it.

Lord Kelvin - Scientist

Chapter Summary

The previous chapter has presented how generic quality indicators could be retrieved from results of CPAM. This chapter uses these results to improve the land-product reliability by optimising these reliability indexes over the land-product model's coefficients. After explaining the underlying model and limitations, comments and recommendations on suitable optimisation methods are made. Finally this framework is applied to an example.

7.1 Introduction

Chapters 4 and 6 have explained how overlapping parts of hyperspectral flight-lines could be used to retrieve an assessment of the reproducibility of a mapping product. Given a specific case-study, getting such a rating leads in turn to several practical questions.

- Is the achieved quality enough for our application?
- Is it the best that could be achieved?

As seen in sections 3 and 6.1.1, the first question can only be answered by the end-user of the results. However, when it is answered with a negative answer, it is quite likely that the data provider will be asked: '*Is it possible to improve this result?*', which is, in essence, similar to the second question.

Since quality assessment as presented in sections 4 and 6 can be performed on a regular basis, it can theoretically be used to tune the inputs used in computations in order to get an optimum result. These inputs are threefold:

- The input data itself - e.g. the hyperspectral data, and some ground-truth: D
- The parameters used for running the models: λ
- The models used for processing the data: $\mathcal{M} : \lambda, D \rightarrow \mathcal{M}_\lambda(D) = R$

Formally, if $\epsilon(R)$ denotes the quality criterion - typically AD or Krippendorff's α - associated to the result R, the goal becomes about finding the inputs minimising the quality criterion:

$$\arg \min_{M,\lambda,D} \epsilon\left[M_\lambda(D)\right] \tag{7.1.1}$$

Section 7.2 summarises and stresses the degrees of freedom available to improve the data. The retained methodology is then summarised in section 7.3 before being applied to an illustrative application in section 7.5.

7.2 Framework & Constraints

7.2.1 Constraints

To start with, some constraints on equation 7.1.1 have to be mentioned. First of all, since the quality indicators derived on chapter 6 are based empirically on the data, they can only be computed once the data has been acquired and pre-processed.

A direct corollary, given the practical cost of a flight survey in terms of time and money, is the fact that the 'raw data' cannot be changed. Redoing the same survey several times is - in almost all cases - not an option. From this point, it might seem that end-users are doomed to live with the retrieved quality.

Looking back into the processing chain, there are however two remaining degrees of freedom: the models and their parameters. Assuming that the processing was performed using state of the art methodologies, it seems quite unlikely that the models could be improved or even changed on the short term. For example, many models are actually based on physical considerations. Swapping them would therefore make very little sense except for a few specific cases which are usually well identified.

There might of course be some specific cases where several alternative models are available, for example different classification methodologies. Even in such a case there would usually only be a handful of available methodologies. One could therefore perform the analysis for each of them, then compare each quality criterion and their confidence intervals to retrieve the most suitable one [Ehrler, 2014].

In general, the only 'true' degree of freedom which remains is therefore the set of parameters used for processing models. Section 7.2.2 reviews the details of their set-up and origins.

7.2.2 Model's Parameters: Origins and Analysis

Before trying to adjust the model's parameters, one should pay attention to their origins, and therefore to the extent they can be tuned. One can indeed identify several cases. Although these categories do not correspond to strict boundaries, they could be divided in the following rough types.

- Empirical parameters set by the users. They can be tuned without limitations, provided one stays within their definition domain. This type does for example encompasses the 'proportion parameter' in most of the applications where results are based on a weighted sum of a 'prior' term - including information about expected behaviour - and a 'data' term accounting for what is actually observed in the data [Rochery et al., 2006; Descamps et al., 2008]. For example an asphalt road can be defined by the prior statement '*its shape should be close to linear*' and the data term: '*it has the spectral response of asphalt*'. The relative importance of these two statements is a model parameter not a priori known.

- Measured parameters. These parameters should be measured along with a confidence interval. It is possible, provided they are not of purely statistical nature, to tune them but this has to be done in conjunction with their measurement uncertainties. This category for example encompasses calibration parameters [Baumgartner et al., 2012; Lenhard, 2012].
- Physical parameters. They are actually fixed by underlying physical laws and cannot be directly tuned. This is for example the case of the absorptions bands induced by a specific mineral [Bayer et al., 2012].

It should be stressed again that these categories are not always strict ones: most parameters are in practice more likely to be a mixture of these. For example, there is a relationship between an amount of atmospheric water vapour and the corresponding absorption. The absorption is however quite likely to be a measured parameter, thus not strictly defining the water vapour content. From a qualitative point of view, one would prefer to optimise parameters 'close to' being empirical, while physical parameters should be left aside.

7.3 Methodology: Improving Results

As seen in section 7.2.1, the original model from equation 7.1.1 can already be restricted to equation 7.3.1, where λ is the set of 'tunable' parameters, while λ_0 is the set of physical, non-tunable parameters:

$$\arg\min_{\lambda} \epsilon \left[M_{\lambda_0,\lambda} \left(D \right) \right] \tag{7.3.1}$$

Even in this simplified case, several points remain to be discussed. Section 7.3.1 reviews some further limitations to the choice of the parameters to be tuned. Section 7.3.2 and 7.3.3, in turn, reviews some limitations of this optimisation process as well as workarounds.

In the following, the set of non-tuned parameters are omitted from the notations and is considered as part of the data.

7.3.1 Which 'Empirical' Parameters?

Before going any further, a first issue has to be addressed. In section 3.3.1, the Monte-Carlo error propagation framework, was dealing with a similar issue: given a probability density function of the input data and parameters, what would the output probability density function look like? The retained solution consisted in randomly sample tuplets of these inputs before propagating them into the system in order to obtain the output distribution. One of the main limitations of the approach was its computational complexity. Given P parameters, the corresponding simulation would involve getting a dense parsing of a P-dimensional space which yields an exponential complexity.

When tuning the parameters to obtain the best achievable result, the search is about looking for a minimum on a P-dimensional space, which is having an exponential complexity in P too. It is therefore leading, in almost all cases to non feasible computation times. A selection of a subset of the parameters should therefore be done.

A second issue is the retrieval of the matches: as seen in chapter 4, the algorithm is actually based on $L2$ data. This means that changing the parameters used for atmospheric or ortho-rectification parameters would potentially lead to a new set of matches, hence preventing an optimisation to take place on them. Optimising the L2-data by comparing pairs of matches is therefore a very complex - if not impossible - task. The parameter tuning therefore only takes place for processing steps of a higher level than $L2$: namely $L3$.

Both arguments lead to the conclusion that the optimisation should take place on land-product generation models only, hence limiting the amount of parameters to be optimised while maintaining the validity of the approach.

7.3.2 Optimisation Constraints

A further issue to be considered is the parameter definition domain. Since the optimisation aims at decreasing the amount of discrepancies between both flight-lines, there is a risk to over tune the parameters to produce an artificial result. For example, a classification might be optimised to allocate to each pixel the same class - independently of its spectral content. The similarity between each flight-line would then be extremely high, although its intrinsic validity would be null and void.

This implies, in practice, that the parameters have to be kept in a plausibility range. For measured parameters, this should be comparably straightforward. On the other hand, empirical parameters might require some analysis before their admissibility domain can be derived. This leads to nothing more than a set of constraints on the parameters' domain.

Alternatively or in conjunction with the aforementioned constraints, a few external constraints might of course be added. These could for example be some 'ground-truth' information such as '*this specific pixel has to be from class 'C'*'.

7.3.3 Over-fitting

A potential issue has finally to be addressed. Depending on the model, the optimisation result might yield a perfect agreement while that the agreement is completely out of bounds when outside of the set of control points. A well-known example of such phenomenon is the so-called Runge effect when performing an interpolation by means of polynomials: even if a polynomial - e.g. a Lagrange interpolation polynomial[1] - is exactly fitting on every control points, it might still be extremely different from the actual data (see figure 7.1).

In order to account for this potential issue, the set of control points has to be split into two subsets. The first one is used solely to retrieve the optimum parameters, the second one being solely used to control the results and ensure that the minimisation process went well. If it has not, some further constraints - depending on the model - would have to be included. From a more generic point of view, the use of a test subset unused for computations is a classical validation method in machine-learning techniques.

The selection of the validation subset is here achieved using the so-called random sub-sampling validation. Due to the large number of retrieved matches, exhaustive cross-validation is indeed not feasible. On the other hand, taking a few percent of the retrieved matches as a validation group should encompasses enough points to grant a statistically valid representation of the complete set.

[1] For Lagrange polynomials, see appendix A.4.5.

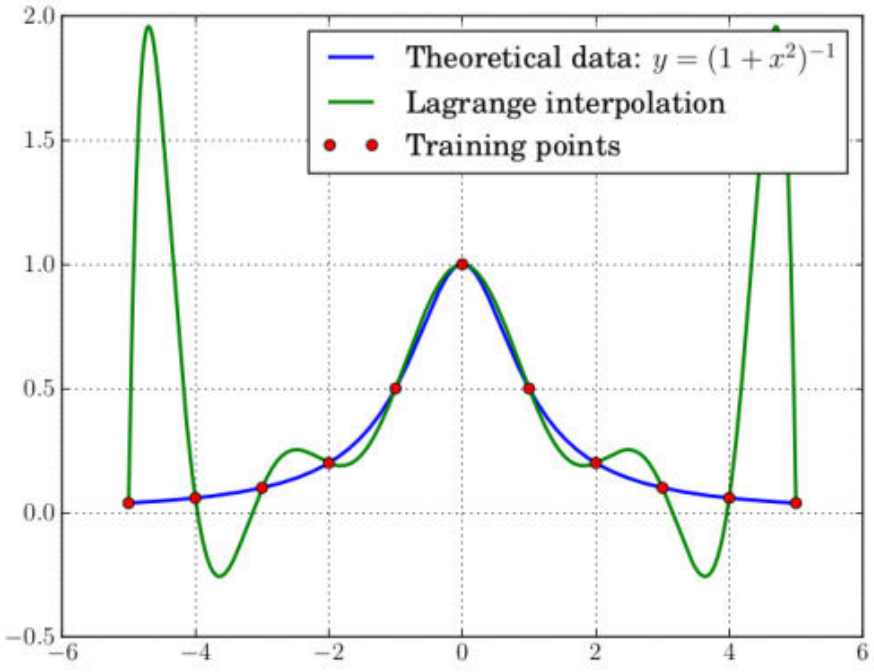

Figure 7.1: Example of over-fitting: when provided with too many control points, a Lagrange interpolation might be subject to the Runge effect: although perfectly fitting to the control points, the interpolation has little to do with the original data.

7.4 Optimisation Method

7.4.1 Generic Considerations

In general, very little is know about the behaviour of the error term with respect to the model's parameters. It might be a non-convex relationship or a non-derivative function. There is furthermore no information about potential local minima. The selected methodology however has to be able to retrieve the global minimum. In order to achieve such goals, the optimisation has to rely on stochastic methods, whose two main classes are simulated annealing and genetic methodologies [Allaire, 2005].

The simulated annealing method was introduced by [Kirkpatrick et al., 1983] and [Černý, 1985]. It consists in randomly exploring the parameter space by moving to a new set of parameters if:

- it leads to a higher reliability.
- or is randomly accepted using an always decreasing probability.

Genetic algorithms on the other hand generate a population of possible sets parameters. It then mutates and mixes these sets iteratively, keeping them with a probability that increases with the reliability they induce, thus mimicking genetic evolution.

The comparison of simulated annealing and genetic algorithm methods have lead to numerous publications about their relative speed and approximations of the optimum. Ross and Corne

[1995] presented empirical experiments on the timetable allocation problem with the outcome that simulated annealing performed better than genetic algorithms. This result was backed up with experiments made on graph optimisation problems which indicated that simulated annealing was both faster and more accurate than genetic algorithms [Lahtinen et al., 1996].

On the other hand, Manikas and Cain [1996] applied both approaches on circuit design issues and obtained the opposite conclusion in terms of accuracy. They however did not mention execution times. Finally, Hosny [2011]; Garg [2009] empirically retrieved similar accuracies for both methods in the fields of pick-up/delivery and crypto-analysis. Hosny [2011] indicates an advantage for the simulated annealing in terms of execution speed. [Arostegui Jr. et al., 2006] indicated similar findings in the field of facilities location: depending on the case study, simulated annealing and genetic algorithms are alternatively better than the other one.

From these studies, it appears that the simulated annealing seems to be at least equivalent to the genetic algorithm when taking into account both execution times and final accuracies. Since it also requires less tuning than the genetic algorithms (for example, no cross-over of the population), simulated annealing is preferred over the latter. This should not hide the fact that there is no intrinsic hierarchy between both methods.

7.4.2 Other Methodologies

It should be stressed again, that, as soon as the parameters have a nicer relationship to the reliability function, other optimisations methodologies - usually much more efficient - should be preferred. The potential variety of specific cases is however too wide to be addressed exhaustively in the general case. As potential particular cases, one could cite [Allaire, 2005; Weise, 2009]:

- If $\epsilon_D(\lambda)$ is convex and differentiable, gradient descent methods could be used.
- Linear programming methods could be used when the reliability function and associated constraints can be expressed as a linear combination of the model's parameters: $\epsilon_D(\lambda_1, \ldots, \lambda_k) = a_1\lambda_1 + \ldots + a_k\lambda_k$.
- Local optimisation methods could be used when a solution 'near' the global optimum is known.

When ϵ_D is derivable, direct retrieval of the optimum parameters might be obtained by computing the λ meeting the optimality conditions:

$$\begin{cases} \dfrac{\partial \epsilon_D}{\partial \lambda}(\lambda) & = & 0 \\ \dfrac{\partial^2 \epsilon_D}{\partial \lambda^2}(\lambda) & \geqslant & 0 \end{cases} \tag{7.4.1}$$

Except for trivial models, these cases are however quite unlikely to occur and one has to rely on simulated annealing.

7.4.3 Simulated Annealing

Now that the simulated annealing has been proposed as the standard optimisation method for this work, its background is reviewed in more detail. The next paragraphs presents the details of the algorithm before reviewing its convergence.

Algorithm

The main idea of the algorithm is inspired by the annealing method used in metallurgy [Kirkpatrick et al., 1983]: a material is slowly cooled in order to allow its internal crystals to grow to an optimum size, ensuring better mechanical properties. In simulated annealing this slow decrease of the system's energy is translated into an iterative modification of the system - here the parameters to be tuned - along with a slow decrease of the acceptance rate for getting into a 'worse state that the current one'. The possibility to decrease the quality of the current parameters-set enables to avoid been trapped in local minima, thus getting an optimum.

In practice, each configuration - p - of the material leads to a specific energy - E. The probability for the material in configuration p, when at thermal equilibrium at temperature is T, can be expressed as being proportional to the Boltzmann distribution (also called Gibbs distribution):

$$P(p) \propto \exp\left(-\frac{E(p)}{k_B T}\right) \tag{7.4.2}$$

k_B being the Boltzmann's constant. This yields that the ratio of probabilities corresponding to two states is given by the so-called Boltzmann factor, depending solely on each state energy and on the temperature.

$$\frac{P(p_1)}{P(p_2)} = \exp\left(-\frac{E(p_1) - E(p_2)}{k_B T}\right) \tag{7.4.3}$$

This property can then be linked to the Metropolis-Hasting algorithm to obtain random samples of the configuration according to the corresponding distribution at each given temperature [Hastings, 1970]. For high temperatures, the actual energy for each configuration plays little role, thus allowing to explore their whole set without being stuck at a local minimum. At low temperatures, the configurations corresponding to lower energy is preferred, thus ensuring a convergence to the minimum.

In the following, the material configuration - or state - is simply the set parameters to be tuned λ, the energy corresponding to the score of the AD coefficient obtained for λ, given a specific data-set D: $AD(\lambda)$. The formal method is given in algorithm 4. The proposed implementation corresponds to a piecewise constant geometrical cooling [Brooks et al., 2003; Ortner et al., 2007]. Other equivalent alternatives however exist but lead to greater computation times [Ortner et al., 2007; Weise, 2009].

As stated, there are fives parameters for the algorithms to be determined.

- The initial state: in practice this does not play a role, provided the initial temperature is high enough to allow an almost random parsing of the initial configurations.
- The decrease parameter, c, it is typically set to 0.995.
- The start and end temperatures T_s and T_e as well as the step length m are set empirically (see section 7.5).

Convergence

The previous sections have presented the simulated annealing as a method which allows to obtain the minimum of a mapping[2] with almost no external hypothesis. This strength however has a cost: the computational complexity of the algorithm.

[2]Here 'mapping' is to be understood in its mathematical sense.

Algorithm 4: Simulated Annealing

Data: λ_0: an initial state
Data: T_s: an initial temperature
Data: T_e: a final temperature, close to $0K$
Data: c: the decrease rate for the temperature: $0 < c < 1$
Data: m: the iso-temperature step length

Result: λ_i: optimised set of parameters

Initialisation:
$T_i = T_s$
$\lambda_i := \lambda_0$
$E_i = AD(\lambda_i)$
while $T_i > T_e$ **do**
  for k *from 1 to* m **do**
    Draw a new set of parameters λ_{new} close to λ_i
    $E_{new} = AD(\lambda_{new})$
    `// Compute` λ_{new}`'s reliability`
    Draw a random number p
    if $p < \exp[-(E_{new} - E_i)/T_i]$ **then**
      $\lambda_i := \lambda_{new}$
      $E_i := E_{new}$
    end
  end
  `// Update temperature:`
  $T_i := cT_i$
end

Various proofs of the algorithm convergence to a global minimum as well as for its sub-cases have been proposed [Anily and Federgruen, 1987; Gidas, 1985; Granville et al., 1994; Hajek, 1988; Ingber, 1993; Nolte and Schrader, 2000] etc. As stressed by [Nolte and Schrader, 2000; Weise, 2009], these proofs are however theoretical: the convergence rate for finite sets of configurations is slower than a systematic enumeration of all configurations. Empirical results however show that the algorithm converges faster to good approximations of the global optimum although no proof exists in the general use case [Nolte and Schrader, 2000; Weise, 2009]. A workaround for this drawback consists of running several instances of the algorithm, and comparing the end results: thus checking for their consistency.

In practice, given a start and an end temperatures, T_s and T_e, the choice of a geometric decrease of the temperature by a factor c (i.e. $T_{i+1} = cT_i$) yields a total number of iterations, I, equals to:

$$c^I T_s = T_e \Rightarrow I = \frac{\log(T_e) - \log(T_s)}{\log(c)} \tag{7.4.4}$$

7.5 Application

The previous sections have proposed a generic and quite abstract methodology that can be used to compute the mapping model's parameters that would lead to an optimal reliability. This chapter applies the methodology to a case study: namely the classification applied to the Sokolov 2010 data-set (see section 6.1.2). The optimisation focuses on the overlap between flight-lines 2 and 3 where most of the mining activities are taking place: performing it over the complete scene, thus including a higher percentage of vegetation pixels, would work in exactly the same way, but would produce slightly different results, paying more attention to vegetated classes.

To start with, section 7.5.1 reviews the details of the classification approach, and the potential parameters to be tuned. Section 7.5.2 reviews the simulated annealing approach used for this specific case. Finally section 7.5.3 presents a summary of the results.

7.5.1 Classification Model

Background

As already summarised in section 6.1.2, the classification was performed on 59 end-members. 'Dark pixels' from the imagery, containing too low reflectance values, were first masked out. Each pixel was then compared by means of a weighted spectral angle to each of the end-member classes, the closest one being considered as being the 'right' class. By weighted spectral angle, it is meant a weighted sum of the spectral angles obtained for the visible part of the spectra, the near-infra-red part and the short-wave infra-red part.

In this general form, the classification approach yields a total of 8 parameters:

- The 'too-low-reflectance-value' threshold. Since it is a depending directly on the sensor signal-to-noise ratio, it was left out of the optimisation.
- The choice of the spectral metric: here the spectral angle. This actually is an empirical parameter. However, section 4.3.1 leads to the conclusion that a change of spectral metric does not bring any gain. This parameter is therefore left aside too.
- The three weighting coefficients: these are purely empirical parameters and can therefore be tuned.
- The limits defining the weighted domain: in practice, this is the index of two bands defining the limits between VIS and NIR and between NIR and SWIR. These are, in general, also empirical and can therefore be tuned.
- The number of sub-portions in each spectra should be subdivided in (here: 3). Again this is a quite empirical decision, its only justification being the fact that minerals are mostly identified by features in the SWIR part of the spectrum, vegetation influencing more the NIR part. Allowing a variability of this parameter would however drastically increase the dimensionality and therefore the search space for the weighting domains and coefficients. It was therefore left aside.

Modelling

The considerations presented in the previous paragraphs yield a total of five parameters that could potentially be optimised:

- The last band of the first part of the spectra: b_1.
- The first band of the last part of the spectra: b_2.
- The weighted coefficient for the spectral angle mapper applied to the first part of the spectra: λ_1.
- The weighted coefficient for the spectral angle mapper applied to the second part of the spectra: λ_2.
- The weighted coefficient for the spectral angle mapper applied to the third part of the spectra: λ_3.

A direct constraint that could be used to limit the search space is to demand that the weighting coefficients are positive and normalised, each coefficient being greater than 0.1.

$$\lambda_1 + \lambda_2 + \lambda_3 = 1 \tag{7.5.1}$$

The former constraint implies that at least one part of the spectra has an influence on the classification. In order to complement this a final requirement insures that each domain of the spectra is non empty and wide enough to imply an influence on the classification. A further constraint was therefore added, imposing that there are at least 20 bands in each spectral domain (see equation 7.5.2, B being the total number of bands).

$$\begin{cases} b_1 & \geqslant & 20 \\ b_2 - b_1 & \geqslant & 20 \\ b_2 + 20 & \leqslant & B \end{cases} \tag{7.5.2}$$

These constraints enforce that the optimisation leads to plausible results. There is therefore no need for any further constraint. Finally, 10% of the retrieved matches were randomly selected to be left aside from the computations, and were used as a validation subset.

7.5.2 Obtaining Optimisation Parameters

Once the problem has been set up as explained in section 7.5.1, the optimisation task should be started. As already stated in section 7.4.3, the simulated annealing parameters have to be set up before being able to run the optimisation.

The first annealing parameter to be searched was the start temperature: T_s. In order to be able to browse the whole set of configurations, it has to be high enough to balance the actual reliability. On the other hand, if it is set to a too high value, this yields useless computations. In order to obtain an estimate, step 6 from algorithm 4 was performed for large m and several temperatures. For each step, the evolution of the parameters was recorded and the smallest temperature allowing a seemingly random configuration evolution was chosen as being $T_s = 20$.

The end temperature, T_e, was set to be close to 0: $T_e = 0.045$ as it was observed that little or no changes were occurring for smaller temperatures. The decrease rate of the temperature was arbitrarily set to be $c = 0.995$. The initial state was randomly set - due to the sufficiently high temperature chosen for T_s, as well as 1000 initial blank iterations to ensure proper initialisation of the system, this does however not play a role.

Finally, the iso-temperature step lengths were also obtained empirically. Step 6 from algorithm 4 was once again run for several temperatures within T_e and T_s, m being set to the maximum number of iterations required to obtain a thermodynamic equilibrium: $m = 50$.

One can note that this implies a total of mI iterations:

$$mI = m\frac{\log(T_e) - \log(T_s)}{\log(c)} = 50 * \frac{\log(0.045) - \log(20)}{\log(0.995)} \approx 61000$$

Accounting for about half a second per iteration, this yields a computation time of about 8 : 30 hours per run. Finally, the optimisation was run 10 times to ensure the global convergence, which yielded a total computation time of about 3 days and a half.

7.5.3 Results and Analysis

All optimisation results are displayed on table 7.1. The optimisation evolution for a typical case - run #6 - is displayed on figure 7.2, while the evolutions for the remaining executions are displayed on appendix B.2. Please note that scores are represented in a nominal data fashion with p_0, although computations were made on the more general but slightly less intuitive AD.

Run	b_1	b_2	λ_1	λ_2	λ_3	$1 - p_o$	$1 - p_o$ validation
1	40	61	0.1143	0.7856	0.1001	11.29%	15.34%
2	41	61	0.1156	0.7854	0.1000	10.80%	14.72%
3	41	61	0.1222	0.7774	0.1004	11.05%	14.07%
4	38	69	0.2284	0.6641	0.1075	22.82%	27.84%
5	38	69	0.2211	0.6789	0.1000	22.79%	28.45%
6	41	61	0.1272	0.7728	0.1000	10.83%	13.15%
7	38	69	0.2153	0.6767	0.1080	22.80%	28.78%
8	38	69	0.2358	0.6492	0.1150	22.67%	29.06%
9	41	61	0.1059	0.7941	0.1000	10.87%	14.69%
10	41	61	0.1128	0.7889	0.1000	10.79%	15.64%

Table 7.1: Optimisation results for all the ten runs.

To start with, it can be noticed that the actual reliability optimum was achieved for $\lambda_1 \approx 0.12$, $\lambda_2 \approx 0.78$, $\lambda_3 \approx 0.11$, $b_1 = 41$ and $b_2 = 61$, yielding a final p_0 of about 86% (validation score).

As it can be seen on table 7.1, some runs did not completely converge and stayed stuck at a local minima (e.g. runs 4 or 5, figures B.4, B.5). The other runs did converge to around $b_1 = 41$, $b_2 = 61$, $(\lambda_1, \lambda_2, \lambda_3) = (0.12, 0.78, 0.10)$. Although it is definitively not the exact global optimum, any further optimisation seems futile since the discrepancies of p_0 between the computations and the validation subsets are much greater than the variations observed for p_0. Any further optimisation would therefore be 'lost in over-fitting' and the aforementioned result is considered as the optimum.

A further interesting feature is the brutal score jump observed for the runs which did indeed converge to the global optimum. It likely corresponds to an energy pit corresponding to a narrow parameter domain. This specificity is likely the reason why some non-optimum runs ignored it: simply because no set of parameters in the right-domain were drawn. For example on the fourth run (see figure B.4), a 'right' parameter set was drawn around the 20000^{th} iteration before being rejected due to the still too high temperature. At a lower temperature, the annealing did not come back at these points of the parameter space. Although this is theoretically an indicator of

a 'too quick' temperature decrease, it is in practice only a practical example of the limitations explained in the paragraph about convergence.

It can furthermore be observed that small perturbations of the three weighting coefficients have little or no impact on the final score although the band selection is important. One can furthermore note the low score obtained for the SWIR-coefficient - λ_3. It can easily be explained by the higher amount of noise with respect to the NIR and VIS domain. Since the spectral angle is sensitive to noise [Robila, 2004], it is logical that the SWIR domain is given less credit than its counterpart at low wavelengths.

The score obtained for the validation set should finally be compared to the raw scores obtained without parameter optimisation in section 6.2.4: the original p_0 for all classes was 69%, the one for mineral classes only being about 86%. After the optimisation of the parameters, the final p_0 is also about 86% for not only the mineral classes but also the vegetation ones[3], yielding an increase of about 17%. Despite being only valid for this specific example, this improvement indicates the potential of such a parameter optimisation methodology.

[3]Based on the validation score only.

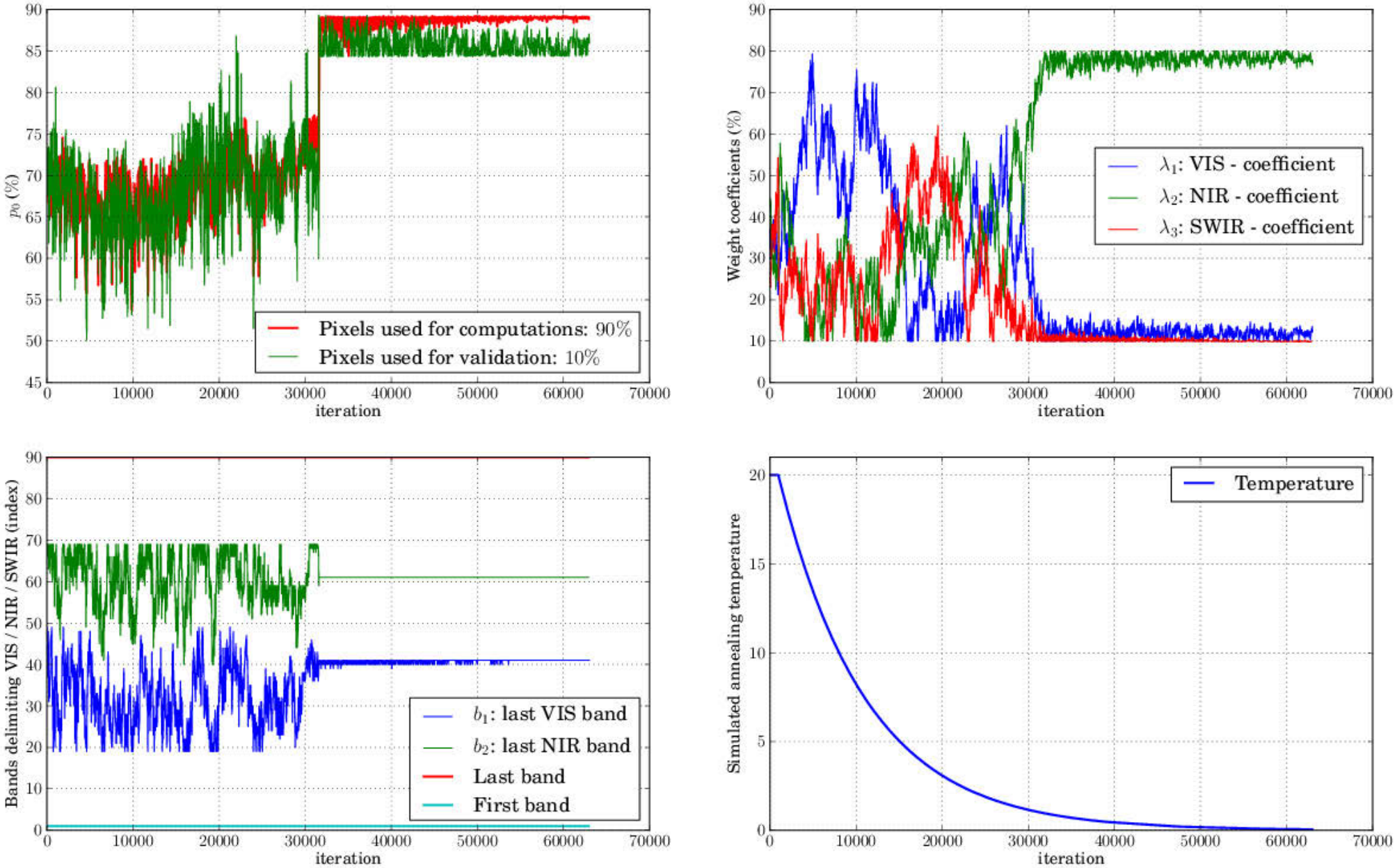

Figure 7.2: Simulated annealing, evolution of the sixth run.

7.6 Conclusions

This chapter has presented a methodology allowing for an improvement of mapping results, provided they were obtained by parameter dependent models. It should be noted that this remains a quite slow method, which moreover scales very badly when the number of parameters increase: each new parameter adds a new dimension to the search space, thus leading to an exponential complexity in terms of the number of parameters.

A further drawback consists of the optimisation method by itself. Since it works on the models parameters, it would not be usable on parameter-free models. For example, it cannot be used on a pure NDVI computation. On the other hand, it could be used to optimise a threshold on NDVI-scores, for example to determine which minimum NDVI corresponds to vegetation and which does not. Besides, the parameters leading to the actual data-set cannot be tuned by these means: this therefore addresses only the land-product generation step.

The method nevertheless allows for an automatic optimisation, thus possibly allowing for an improvement of the products' quality. In the presented example, the initial p_0 was increased from 69% up to 86%. This might be especially useful when dealing with non-linear models, which can hardly be optimised by hand.

Readers might wonder why the retained optimisation took place on AD. Chapter 6 indeed indicated that Krippendorff's α coefficient was actually a better and more acute alternative. The reason for this choice is actually lying in the corresponding complexity. Computing AD is comparably fast - $\mathcal{O}(N)$ - and has almost no memory requirement. Krippendorff's α on the other hand requires a much higher number of operations and memory allocations (see section 6.4.4). Since the simulated annealing optimisation is in turn requiring a huge amount of iterations, it does make more sense to rely on quick-to-compute though slightly less advanced coefficients. A further justification lies in the criticisms expressed against AD: its main drawback was its lack of invariance with respect to the number of evaluation points. However, in the case of a simulated annealing approach, the amount of evaluation points remains the same, thus discarding this drawback and allowing a direct comparison of different AD.

Chapter 8

In Depth Analysis of the Variability

> Some people are always critical of vague statements. I tend rather to be critical of precise statements; they are the only ones which can correctly be labeled 'wrong'.
>
> Raymond Smullyan - Logician

Chapter Summary

This chapter presents how results derived from CPAM in the specific case of two overlapping flight-lines can be used to derive an estimation of the variability of the reflectance. A first section introduces the methodology before conducting in-depth analysis in a second step. Theoretical background about the modelling is then summarised before proposing a algorithm suitable to estimating individual variability. In a final part, it is applied to illustrative data.

8.1 Introduction

Chapter 6 has shown how pairs of pixels as obtained by means of the CPAM algorithm could be used to evaluate mapping results. Similarly, chapter 7 presented a methodology to improve these mapping results given a data-set. However, none of these chapters have provided a fully generic quality assessment.

To start with, one can notice that the achieved reliability assessment cannot be linked at all with any issue from the pre-processing chain. It is in particular absolutely impossible to retrieve any information about bottle-necks limiting the accuracy of a Krippendorff's α coefficient. Although this goal seems to be extremely difficult to achieve, a first step would be to gain a better understanding of the reflectance variability.

A second illustration can be the model's parameter optimisation steps, as presented in chapter 7. The proposed method is clearly only valid for a specific data-set: even if the method is generic enough to be run on virtually every problem, this has to be done on each new issue, hence requiring a very high computational load over time. An alternative approach would be to analyse the variability present in the $L2$ data and try to reduce it directly, thus directly providing

the 'best' $L2$ data, instead of trying to adapt the $L3$ mapping to faulty reflectance. This could however only be performed once this $L2$ radiometric variability would become assessable.

Finally, indicators presented in chapter 6 can be computed on each new result, but do not allow any prior guess about another new mapping application result. Gaining an understanding of the radiometric variability present at the $L2$ level would, again, open the way for an a priori assessment of this variability.

This chapter presents elements aiming at bridging this gap. To start with, section 8.2 lists the underlying theoretical principles of the proposed approach. Section 8.3 then reviews the information that could be retrieved about the data: these results are then linked to some theoretical considerations in section 8.4. Finally, section 8.5 presents an approach for estimating the radiometric variability associated with spectral bands before applying them to an illustrative data-set in section 8.6.

8.2 Retained Modelling

Unless specified otherwise, this chapter only deals with pairs of flight-lines sharing the same setup and in particular the same viewing geometry. Notions presented here do furthermore strictly apply on $L2$ data, i.e. on reflectance.

Section 8.2.1 starts by stating the problem and proposing a modelling while section 8.2.2 reviews existing approaches for solving it. Finally section 8.2.3 deals with some limitations of the approach.

8.2.1 Principles

The following is based on the results returned by the CPAM method and would only apply to the case of parallel flight-lines sharing the same viewpoint. Under these conditions, it is widely accepted that the observed noise is, at least as a first approximation, of additive nature [Atkinson et al., 2003; Chang and Du, 2004; Farzam and Beheshti, 2011; Green et al., 1988; Letexier and Bourennane, 2008; Parra et al., 2000; Sagar et al., 2014]. It follows that each pixel signal - s - can be decomposed as the 'true/real' reflectance - r - plus some noise or variability - v. In the following, the noise is considered to also encompass the image variability caused by slight scene changes between two consecutive acquisitions - e.g. changes of pixels' footprints, which could also be modelled by additive effects induced by change of the pixels' end-members [Rogge et al., 2006].

When disposing of CPAM outputs there are not one, but two observations of the variability available, even though there are three unknowns: the underlying purely theoretical 'truth' r, and two types of 'variability' associated with it. In order to get rid of the actual 'real' term, r - the difference between both observations are computed:

$$\begin{aligned} o - \tilde{o} &= (r + v) - (\tilde{r} + \tilde{v}) \\ &= (r + v) - (r + \tilde{v}) \\ &= v - \tilde{v} \end{aligned} \tag{8.2.1}$$

Since o and $\tilde{o}$ are known - they are the pixel's values - an estimation of $v - \tilde{v}$ is directly available. By constructing its histogram, it is furthermore possible to gain an estimate of the probability density function corresponding to this difference: f_d - computed for each spectral channel.

Noting f_v (resp. $\tilde{f}_v$) the probability density function from which v (resp. $\tilde{v}$) is drawn and writing $\tilde{f}_v^-$ the reversed distribution of $\tilde{f}_v$ - $\tilde{f}_v^-(x) : x \to \tilde{f}_v(-x)$ - this yields the convolution relationship displayed on equation 8.2.2.

$$\begin{aligned} f_d(x) &= \int_{-\infty}^{\infty} f_v(t)\, \tilde{f}_v^-(x-t)dt \\ f_d &\equiv f_v \star \tilde{f}_v^- \end{aligned} \tag{8.2.2}$$

Equation 8.2.2 however implies a few limitations. First of all, the equivalence is only an approximation. The histograms are indeed only approximations of the actual probability density functions. f_d is furthermore computed on the matches retrieved as explained in chapter 4, which includes a few outliers (see section 4.4.3), thus slightly altering the estimation of f_d. Finally a last technical argument has to be investigated.

The sum of two independent random samples is defined by the convolution of the two corresponding density functions. In practice, this independence condition is only met 'on average'. For example, both flight-lines cannot be acquired at exactly the same time: in practice they are acquired a few minutes appart. It can, in turn, occur that over one field, the grass was cut. The pixel's pair retrieval would then likely produce a pixel representing 'cut grass' versus a pixel corresponding to 'normal grass'. The discrepancies between both classes are definitely not independent. It could however be argued that such cases are only affecting very few pixels. Out of these, these discrepancies are likely to impact CPAM results, thus discarding most of the affected pairs. Nevertheless, a small subset might remain in the final set of pairs, slightly affecting f_d.

The final model is therefore the one stated on equation 8.2.3. f_d, the distribution of the differences of the variability is the only parameter known. f_v and $\tilde{f}_v^-$ are both unknowns to be estimated, while η is simply summarising the aforementioned unknown nuisances. This yields a clearly ill-posed problem.

$$f_d \equiv f_v \star \tilde{f}_v^- + \eta \tag{8.2.3}$$

Section 8.2.2 reviews some techniques developed to address this issue.

8.2.2 Existing Approaches

As stated in equation 8.2.3, the problem is basically becoming a so-called blind deconvolution problem which has received a large interest in the field of image processing. Although this implies dealing with bi-dimensional data for f_d, f_v and $\tilde{f}_v$, the problem remains basically the same as in eq. 8.2.3, where f_v would represent the 'true' image, while $\tilde{f}_v^-$ would be analogue to a blurring kernel [Kundur and Hatzinakos, 1996; Starck et al., 2002; Wipf and Zhang, 2013]. In general, the solutions are searched in the form of minimisers of equation 8.2.4.

$$\arg\min_{f_v, \tilde{f}_v^-} \| f_d - f_v \star \tilde{f}_v^- \|_2 \tag{8.2.4}$$

Several authors have tried to raise the ill-posedness of the problem by introducing smoothness constraints on the functions to be retrieved. The method proposed by Tikhonov consists of adding new terms to be minimised. Given two high pass filters H and $\tilde{H}$, and two weighting parameters λ and $\tilde{\lambda}$, this typically leads to minimising equation 8.2.5.

$$\arg\min_{f_v, \tilde{f}_v^-} \| f_d - f_v \star \tilde{f}_v^- \|_2^2 + \lambda \| H(f_v) \|_2^2 + \tilde{\lambda} \| \tilde{H}\left(\tilde{f}_v^-\right) \|_2^2 \tag{8.2.5}$$

Refinements proposed by You and Kaveh [1996] imposed that the H_1-norm of the two convoluted functions should be minimised too. Similarly, Chan and Wong [1998] proposed to retrieve the solution minimising the total variation of the solutions: $\int |\nabla f_v(t)|\, dt + \int \left|\nabla \tilde{f}_v^-(t)\right| dt$. Krishnan

et al. [2011] on the other hand based the retrieval on a minimisation of the ratio from the l^1 and l^2 norms as a measure of sparsity in order to retrieve the 'true sharp' image.

Beside these variational methods, a set of deconvolution algorithms based on maximum likelihood have been proposed for deconvolution by means of Bayesian analysis such as expectation maximisation [Dempster et al., 1977] and maximum likelihood of the retrieved image with respect to the observed data [Kundur and Hatzinakos, 1996; Levin et al., 2009; Pankajakshan et al., 2008]. The best example of a Bayesian methodology is probably the Richardson-Lucy algorithm applied to the case where noise - η - can be modelled by a Poisson process (photon noise) [Richardson, 1972; Lucy, 1974]. Although several different methods and algorithms classes have been proposed (e.g. maximum a posteriori [Shan et al., 2008] or variational Bayesian [Levin et al., 2011]), the ill posedness is always tackled by an introduction of smoothness condition or other arbitrary prior hypothesis on the convoluted terms such as a Gaussianity of the noise [Jalobeanu et al., 2002].

The main drawback of these families of approaches lies in their underlying hypothesis: convoluted terms retrieval is based on smoothness and sharpness arguments - or at least piece-wise smoothness. Although this actually makes sense in the case of natural images affected by a blurring kernel, there is no element that allows foreseeing such a property of the reflectance's variability in a quantitative way.

On the other hand, since in this case f_d, f_v and $\tilde{f}_v$ are probability density functions, other hypotheses can be exhibited. This is done in section 8.3 where an analysis and review of f_d properties takes place.

8.2.3 Differing Viewpoints

Before analysing further the details of f_d, a last point should be discussed: what about flightlines with differing viewpoints, thus including the BRDF effect? Such an effect can be modelled in most cases by a multiplicative factor largely overwhelming the additive noise [Dorigo, 2008; Robila, 2004]. The underlying model from equation 8.2.1 could then be retrieved by working on the logarithm of the reflectance and making use of its limited development:

$$\begin{array}{rcl} \log(o) & = & \log(rv + o\,(rv)) = \log\left[rv\,(1 + o(1))\right] \\ & = & \log(r) + \log(v) + \log\left((1 + o(1))\right) \\ & \approx & \log(r) + \log(v) + o(1) \\ & \approx & \log(r) + \log(v) \end{array} \tag{8.2.6}$$

Although it is possible to perform the computations, this leads - from a physical point of view - to several issues. To start with, the BRDF is dependent of the viewing angles. Retrieving pairs of independent variability $(v, \tilde{v})$ could theoretically be done by spitting pairs of observations into several groups, each group corresponding to a specific view-angle domain.

A second issue to be addressed is the dependency of the BRDF with the land-cover classes [Peltoniemi et al., 2005; Los et al., 2005]. Again one could think about splitting one more time pairs of observations depending on a pre-classification, thus discarding this dependence. This would however imply removing from any further computations all pairs of pixels corresponding to a mixture of land-covers. In practice, most scenes moreover present an overwhelming majority of vegetation pixels. There would then likely be too few pairs of observations corresponding to other classes to perform any valid assessment.

Finally one could think about performing these computations on vegetation pixels only. The BRDF for such classes is however dependent not only on land-cover and acquisition geometry, but also on the physiological status of the vegetation [Hilker et al., 2008]. This makes it almost

impossible to perform such an analysis in presence of BRDF effects. The following therefore focuses on pairs of flight-lines sharing the same acquisition geometry.

8.3 Analysis of Deviations

The first step was to check whether f_d corresponds to a known distribution, from which information about f_v and $\tilde{f}_v$ could be inferred. Section 8.3.1 presents results based on classical distributions, while section 8.3.2 reviews some more advanced results.

8.3.1 Classical Distributions

To start with, classical distributions were fitted on f_d in terms of maximum likelihood: the corresponding details of the estimators are provided in appendix C.2. For some distributions, which were defined over only a subset of the real domain, rescaling and translations were performed on the data to fit the admissible domain for f_d. Being a difference of reflectance, f_d is indeed defined over $[-1; 1]$. Finally and in order to avoid numerical issues for values too close to zero, a rescaling of the observations took place from $[-1; 1]$ to $[-10; 10]$.

Fitted distributions were the Gaussian, the Student's t, the logistic, the Laplace, the Cauchy, the exponential-power, the log-normal, the gamma, the Levy, the Rayleigh, the Beta and the χ^2 distributions. Tests were performed on the only data-set containing perfectly overlapping flight-lines: this is flight-lines $4A$ and $4C$ from the Kaufbeuren 2012 data-set.

As an example, results for bands 10 are presented on figure 8.1. As it can be easily seen, none of these classical distributions could fit efficiently to the data. Results for other bands yield similar issues.

8.3.2 Which Distribution?

The failure of a modelling by means of classical distributions leads to a reconsideration of the problem. According to the modelling explained in equation 8.2.1, f_d *has to* represent a sum of two random distributions. Furthermore, the variability represented by f_v and $\tilde{f}_v$ are likely to be due to a large number of small influences.

If all these influences were independent identical terms with finite variances, the classical central limit theorem states that f_v and $\tilde{f}_v$ should tend to be Gaussian as the number of influences tend to be infinite. In practice, this might however be a quite extreme hypothesis. A generalisation of this classical central limit theorem however exists [Janicki and Weron, 1994; Nolan, 2013]:

Property 1. *Generalised central limit theorem: Given a sequence of identical independently distributed variables Y_n and two sequences of real positive numbers $\{a_n, b_n\}$, then Z_n converges to a stable distribution in law (the reciprocity being true), where:*

$$Z_n = a_n + \left(\sum_{k=1}^{n} Y_k \right) / b_n$$

The exact definition of a stable distribution - also called levy alpha stable (LAS) or paretian distribution - is given in section 8.4.

It should be noted that this is a generalised version of the central limit theorem, hence encompassing it. It does however remove the requirement for having finite variances, hence allowing the use of a heavy tailed distribution. This in turn meets the results presented in [Aiazzi

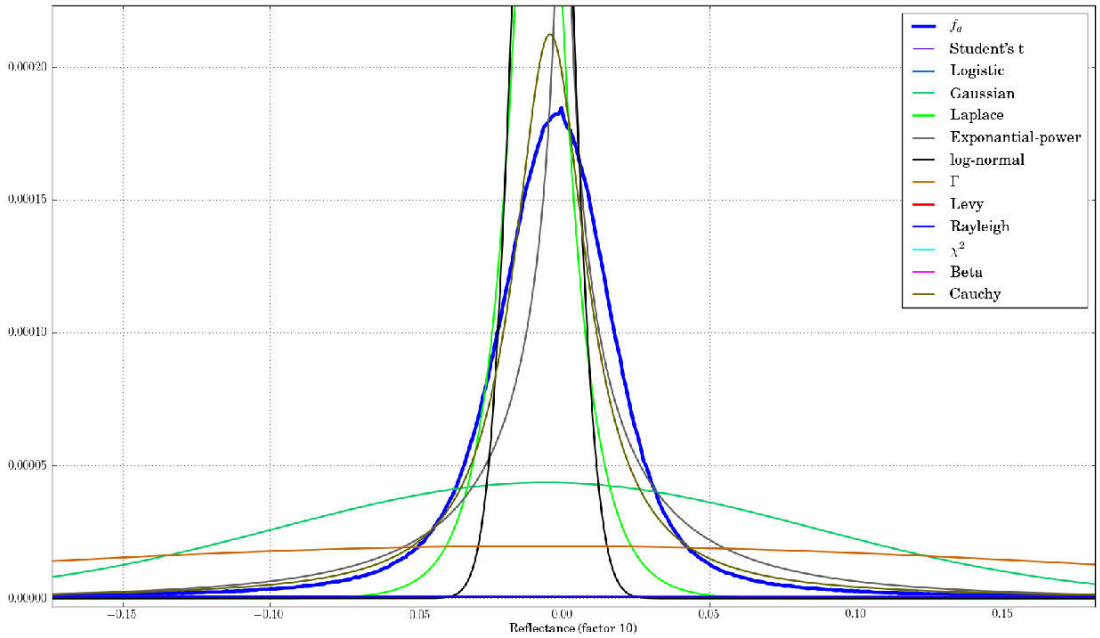

Figure 8.1: Probability density functions obtained by maximum likelihood fitting for classical distributions: example of band 10 at 448 nm (Kaufbeuren 2012, flight-lines 4A and 4C).

et al., 2006] who reported that the noise they observed in hyperspectral data was actually heavy tailed.

In practice, LAS distributions are not a purely theoretical tool and have been widely used for modelling financial data [Mittnik et al., 1999b; Nolan, 2014], noise (of metallic films) [Janicki and Weron, 1994], traffic [Bates and McLaughlin, 2000]. For this work, when fitted to f_d, a stable distribution indeed corresponds to the data: see for example figure 8.2 for band 10; other bands lead to similar results.

Beside these considerations, the LAS distribution presents some other interesting features. For example the sum of two random variables taken from a stable distribution is also a stable distribution, hence providing further determination for f_v and $\tilde{f}_v$ in relationship with f_d. The following section reviews some of the properties of the LAS distribution.

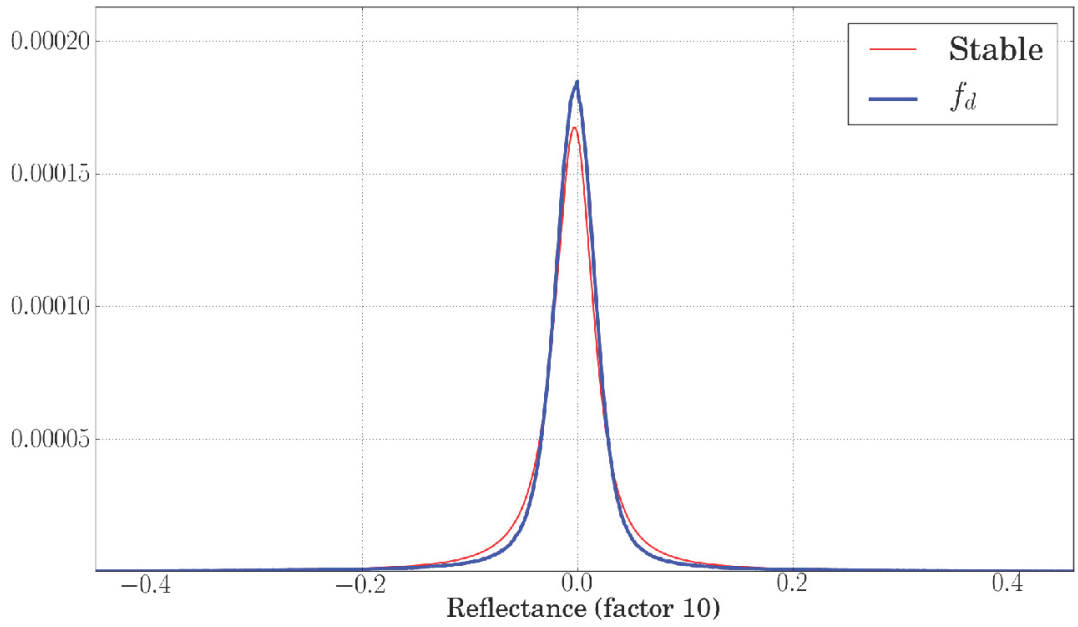

Figure 8.2: Fitting of the stable distribution of f_d: example of band 10 at 448 nm (Kaufbeuren 2012, flight-lines 4A and 4C).

8.4 Stable Distribution

This section reviews some of the main properties of the stable distribution. To start with, section 8.4.1 reviews the notations and parametrisation used. Sections 8.4.2 and 8.4.3 present some of its main properties, while section 8.4.4 reviews how it can be fitted on a data-set.

8.4.1 Notions and Definitions

The main limitation of the stable distribution is the fact that they do not have an analytical representation. They are instead defined by means of their characteristic function as defined in equation 8.4.1.

$$\widetilde{\mathcal{S}_{\alpha,\beta,\gamma,\delta}}(\omega) = \exp\left[-|\gamma\omega|^{\alpha} + i\left(\omega\delta + \beta|\gamma\omega|^{\alpha} sgn(\omega)\Psi_{\alpha}(\omega)\right)\right] \tag{8.4.1}$$

Where sgn is the sign function (see section A.4.7) and Ψ_{α} is defined in equation 8.4.2. When α equals one, $sgn(0)\Psi_1(0)$ is set to zero [Nolan, 2013].

$$\Psi_{\alpha} : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R} \\ \omega & \rightarrow & \begin{cases} \tan(\pi\alpha/2) & \text{if} \quad \alpha \neq 1 \\ -2/\pi \log(|\omega|) & \text{else} \end{cases} \end{cases} \tag{8.4.2}$$

Its probability density function can, by definition, be obtained by applying a Fourier transformation to its characteristic function (see equation 8.4.3). This however does not lead to any analytical, closed form expression and has to be estimated numerically.

$$\mathcal{S}_{\alpha,\beta,\gamma,\delta}(t) = \int_{-\infty}^{\infty} e^{-2i\pi\omega t} \exp\left[-|\gamma\omega|^{\alpha} + i\left(\omega\delta + \beta|\gamma\omega|^{\alpha} sgn(\omega)\Psi_{\alpha}(\omega)\right)\right] d\omega \tag{8.4.3}$$

The parameters can be understood from equation 8.4.1 as being:

- $\alpha \in]0;2]$ rules the 'spread' of the distribution. It is however called the *stability parameter*, due to its role in the distribution properties.
- $\beta \in [-1;1]$ breaks the symmetry around δ and is called *skewness parameter*.
- $\gamma \in]0;\infty[$ is merely just a scaling parameter, hence designed as the *scale parameter*.
- $\delta \in]-\infty;\infty[$, by bringing a phase in the characteristic function, is actually shifting the distribution over the real axis. It is therefore commonly called *location parameter*.

One familiar with stable-distributions might notice that the presented parametrisation corresponds to the S_1 parametrisation as defined in [Nolan, 2013]. It is probably the most widely used form in literature [Bates and McLaughlin, 2000; Kogon and Williams, 1995; Koutrouvelis, 1981; Kuruoğlu, 2001; McCulloch, 1986; Nolan, 1998, 2013] although some other parametrisations have been introduced. The so-called S_0 [Nolan, 1997] (or $\mathcal{M}$-like [Laha, 1989]) parametrisation have also been often used (see equation 8.4.4).

$$\begin{cases} \widetilde{\mathcal{M}}_{\alpha\neq 1,\beta,\gamma,\delta}(\omega) &= \exp\left[-|\gamma\omega|^{\alpha}\left(1 + i\beta\tan\left(\frac{\pi\alpha}{2}\right) sgn(\omega)\left(|\gamma\omega|^{1-\alpha} - 1\right)\right) + i\delta\omega\right] \\ \widetilde{\mathcal{M}}_{\alpha=1,\beta,\gamma,\delta}(\omega) &= \exp\left[-|\gamma\omega|\left(1 + i\beta\frac{2}{\pi} sgn(\omega)\left(\log(|\omega|) + \log(\gamma)\right)\right) + i\delta\omega\right] \end{cases} \tag{8.4.4}$$

The two formulations are of course equivalent, provided the correspondences explained in table 8.1. The main advantage of the latter is however its continuity with respects to its parameters at $\alpha = 1$ [Nolan, 1998]. Indeed, for α in the neighbourhood of 1, using classical developments one retrieves $\mathcal{M}_{\alpha=1}$ as the limit case of the general $\mathcal{M}$ expression (see: appendix C.3.3).

Parameter	S_0 / $\mathcal{M}$	S_1
Stability	α	α
Skewness	β	β
Shape	γ	γ
Localisation	$\delta - \beta\gamma\tan(\pi\alpha/2)$ if $\alpha \neq 1$ $\delta - 2\beta\gamma\log(\gamma)/\pi$ else	δ

Table 8.1: Correspondence between $S_0/\mathcal{M}$ and S_1 parametrisations.

In the following, the $\mathcal{M}$ parametrisation is preferred over the S_1, unless when dealing with authors explicitly referring to the latter.

8.4.2 Sum of Stable Distributions

For this study, one of the main advantages of the stable distribution is its stability with respect to the summation. When summing two random variables from two stable distributions - $(\alpha, \beta_1, \gamma_1, \delta_1)$ and $(\alpha, \beta_2, \gamma_2, \delta_2)$ - with the same stability parameter, the resulting random variable is also a stable distribution with parameters $(\alpha, \beta_0, \gamma_0, \delta_0)$ as stated in equation 8.4.5. This

can be directly checked by multiplying both characteristic functions and associating the result with the distribution defined by $(\alpha_0, \beta_0, \gamma_0, \delta_0)$.

$$\begin{cases} \alpha_0 &= \alpha_1 = \alpha_2 \\ \beta_0 &= \left(\beta_1 \gamma_1^{\alpha_0} + \beta_2 \gamma_2^{\alpha_0}\right) / \gamma_0^{\alpha_0} \\ \gamma_0 &= \left(\gamma_1^{\alpha_0} + \gamma_2^{\alpha_0}\right)^{1/\alpha_0} \\ \delta_0 &= \delta_1 + \delta_2 \end{cases} \tag{8.4.5}$$

In practice, this means that when f_d is approximated by a stable distribution, f_v and $\tilde{f}_v$ could also be represented by stable distributions.

8.4.3 Link with Other Distributions

Finally, one remarkable property of the stable distribution is its ability to encompass the Normal (see section C.2.10) and the Cauchy (see section C.2.2) distributions as some of its special cases (see section C.3.1 for details). When α is set to two, one retrieves a normal distribution with mean $\delta/2\pi$ and a standard deviation of $\gamma/\sqrt{2}\pi$, no matter what value β takes. When α is set to 1 and β to 0, one retrieves the Cauchy distribution of location parameter $\mu = \delta/(2\pi)$ and of shape parameter $\gamma = \alpha/(2/\pi)$.

The fact that it is a generalisation of the Gaussian already is by itself an interesting result. Besides being compliant with the fact that it is a generalisation of the central limit theorem, it does indeed link with the already existing work which modelled noise or variability as being Gaussian [Farzam and Beheshti, 2011; Jalobeanu et al., 2002; Sagar et al., 2014].

The link with the Cauchy distribution is however not as straightforward. It is for example well known that spectral lines under homogeneous broadening in emission spectroscopy have a line-width following a Cauchy distribution [Ott et al., 2013]. Despite being an interesting similarity, there does not seem to be a direct link between radiometric variability and such broadening effects.

One should note the 2π coefficients in the aforementioned equivalences. They are coming from the fact that most of the stable parametrisation consider a Fourier transform defined as $\int e^{-i\nu x} \bullet dx$. On the other hand, this document makes uses of $\int e^{-2i\pi\omega x} \bullet dx$. In order to ensure correspondence with previous citations, the characteristic function definitions 8.4.1 and 8.4.4 were not altered accordingly. This would theoretically imply a scaling, $2\pi\omega = \nu$, factor when comparing $\widetilde{\mathcal{M}}$ with other characteristic functions. However, in this specific case, this corresponds to a rescaling of the localisation and scale parameters by 2π, hence the presence of such factors when comparing these parameters with Normal and Cauchy distributions.

8.4.4 Parameter Estimation

Due to the absence of explicit representations of the probability density function of stable distributions, the parameter estimation step is not as straightforward as for usual distributions. From the seventies up to the nineties, several numerical parameter estimation methods have been developed. They can be divided in three main families [Nolan, 2013]:

- Quantile based methods
- Characteristic based methods
- Maximum likelihood methods applied to approximations of the distribution.

The following only provides information about the characteristic function based methods as it can be considered as optimal [Čížek et al., 2005] and was therefore the only one used within this work. For more information about the other methodologies, please refer to appendix C.3.2.

The characteristic function method were initiated by [Koutrouvelis, 1981] and largely extended by [Kogon and Williams, 1995]. Instead of attempting to fit the stable distribution on the actual distribution, it aims at performing this fitting on the characteristic function. Writing the inverse Fourier transform of f_d, $\check{f}_d(\omega) = \exp[R(\omega) + iI(\omega)]$, one can note that comparing the real part of the logarithm from $\widetilde{\mathcal{M}}$ and assuming equality with $\check{f}_d$ yields equation 8.4.6 which depends only on α and γ:

$$\mathcal{R}\left[\log\left(\widetilde{\mathcal{M}}(\omega)\right)\right] = -|\gamma\omega|^{\alpha} = \mathcal{R}\left[\log\left(\check{f}_d(\omega)\right)\right] \tag{8.4.6}$$

Equation 8.4.6 can be transcribed as a linear relationship by means of the logarithm: $\log(\omega)$.

$$\log\left(-\mathcal{R}\left[\log\left(\check{f}_d(\omega)\right)\right]\right) = \alpha\log(\gamma) + \alpha\log(\omega) \tag{8.4.7}$$

A linear fitting of $\log -\mathcal{R}\left(\log\check{f}_d\right)$ over $\log(\omega)$ then yields estimates of both α and γ. Similar manipulations on the argument of $\widetilde{\mathcal{M}}$ yields to equations 8.4.8 or 8.4.9 - depending on retrieved value for α - which can be solved for β and δ according to least squares fitting.

$$\begin{aligned}\mathcal{I}\left[\log\left(\widetilde{\mathcal{M}_{\alpha\neq 1}}(\omega)\right)\right] &= \omega\left(\delta - \beta\gamma\tan\left(\tfrac{\pi\alpha}{2}\right)\right) + \beta\tan\left(\tfrac{\pi\alpha}{2}\right) sgn(\omega)\,|\gamma\omega|^{\alpha} \\ &= \arg\left[\check{f}_d(\omega)\right]\end{aligned} \tag{8.4.8}$$

$$\mathcal{I}\left[\log\left(\widetilde{\mathcal{M}_{\alpha=1}}(\omega)\right)\right] = \omega\left(\delta - \beta\gamma\frac{2}{\pi}\log\gamma\right) - \omega\log\omega\left(\beta\frac{2}{\pi}\gamma\right) = \arg\left[\check{f}_d(\omega)\right] \tag{8.4.9}$$

According to [Čížek et al., 2005], this kind of estimator is preferable for large sample sizes - which is the case here - where it is both fast and efficient.

It should finally be noted that the method presented here slightly differs from the method proposed by [Kogon and Williams, 1995]. To start with, the input samples here are not rescaled through a pre-estimation of the parameters through the quantile based method [McCulloch, 1986]. As a consequence, the number and values of frequencies - ω - used for both regressions is not arbitrarily set up as in [Kogon and Williams, 1995] but on considerations about the validity of $\log(\omega)$ and $\arg\left[\check{f}_d(\omega)\right]$ in terms of numerical computations.

8.5 Retrieving Individual Variability

From the considerations made in previous sections, an algorithm suitable for estimating f_v and $\tilde{f}_v$ from f_d was developed: section 8.5.1 presents the proposed deconvolution model. Section 8.5.2 reviews the corresponding algorithm, while section 8.5.3 summarises the computation steps.

8.5.1 Retained Model

Since f_d appears to be very similar to a LAS distribution, and since this type of distribution is stable by convolution, f_v and $\tilde{f}_v$ should also be close to being stable convolutions. The smoothness conditions presented in section 8.2.2 could therefore be replaced by 'similarity to a LAS distribution' condition. This yields the deconvolution model presented on equation 8.5.1 where s_0 is the LAS distribution which is the closest to f_d, while s_1 and $\tilde{s}_1$ represent two 'plausible' LAS distributions such as $s_1 \star \tilde{s}_1 = s_0$.

$$\arg\min_{f_v, \tilde{f}_v^-} J\left(f_v, \tilde{f}_v^-\right) = (1-\lambda)\left\|f_d - f_v \star \tilde{f}_v^-\right\|_2^2 + \lambda\left(\|f_v - s_1\|_2^2 + \|\tilde{f}_v^- - \tilde{s}_1\|_2^2\right) \tag{8.5.1}$$

As expressed in equation 8.5.1, the deconvolution problem aims at finding f_v and $\tilde{f}_v^-$ which are close to f_d when convoluted together, while being similar to the LAS s_1 and $\tilde{s}_1$.

Even in its final form as in equation 8.5.1, the deconvolution problem is not completely setup. First of all, one has to account for the weighting parameter - λ - ruling the relative influence of the data-term $\left[f_d - f_v \star \tilde{f}_v^-\right]$ against the prior-term $\left[\|f_v - s_1\|_2^2 + \|\tilde{f}_v^- - \tilde{s}_1\|_2^2\right]$. A logical set-up would be to relate it to the actual similarity between f_d and s_0. When both are actually perfectly matching, one could indeed consider that there is no noise at all, thus that the prior term could be omitted. Taking advantage of the fact that f_d and s_0 are actually probability density functions, hence normalised with respect to the L_1 norm, yields λ to be expressed as in equation 8.5.2.

$$\lambda = \frac{1}{2}\int_{-\infty}^{\infty} |f_d(t) - s_0(t)| dt \tag{8.5.2}$$

The last point to be considered is the choice of s_1 and $\tilde{s}_1$. Since this modelling assumes that both flight-lines should be perfectly overlapping, mostly sharing their acquisition conditions (see section 8.2), the variability between both should be quite similar, hence $f_v \approx \tilde{f}_v$. This in turn yields a further hypothesis on s_1 and $\tilde{s}_1$, namely that both should be similar. Writing α_0, β_0, γ_0 and δ_0 the four parameters defining s_0 and $(\alpha_1, \beta_1, \gamma_1, \delta_1)$ (resp. $\left(\tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\gamma}_1, \tilde{\delta}_1\right)$) the ones defining s_1 (resp. $\tilde{s}_1$), this yields the relation explained in equation 8.5.3.

$$\begin{cases} \alpha_1 &= \alpha_0 \\ \beta_1 &= \beta_0 \\ \gamma_1 &= \gamma_0/2^{1/\alpha_0} \\ \delta_1 &= 0 \end{cases} \quad \text{and} \quad \begin{cases} \tilde{\alpha}_1 &= \alpha_0 \\ \tilde{\beta}_1 &= \beta_0 \\ \tilde{\gamma}_1 &= \gamma_0/2^{1/\alpha_0} \\ \tilde{\delta}_1 &= \delta_0 \end{cases} \tag{8.5.3}$$

It should be noticed that, although the choices of the α and γ parameters are straightforward given the aforementioned arguments and equation 8.4.5, the choice of β and δ parameters is slightly more arbitrary.

The location parameter can indeed be understood as being a measure of the relative bias between both flight-lines. For the sake of simplicity, it was chosen to place it entirely on the $\tilde{\delta}_1$, although other choices would have been just as justified.

Similarly for β, since s_1 should correspond to the reversion of $\tilde{s}_1$, one should theoretically have $\beta_1 = -\tilde{\beta}_1$ which would imply that $\beta_0 = 0$. In practice, β_0 is indeed very close to zero. Since

there is no ground to assume that f_v or $\tilde{f}_v$ should be skewed, they are also assumed to be zero, the slight residual skewness being equally distributed among them.

8.5.2 Getting the Optimum

The previous points having been settled, the next question concerns the optimisation of equation 8.5.1 itself. To start with, it can be remarked that J, although not jointly convex, is convex for both its components taken independently. Such a consideration leads to an alternate minimisation type of solvers, where J is iteratively minimised for f_v, - $\tilde{f}_v^-$ being fixed - before minimising over $\tilde{f}_v^-$, f_v being fixed. Due to this convexity, the minimum of J given $\tilde{f}_v^-$ corresponds to the first order optimality condition:

$$\left.\frac{\partial J}{\partial f_v}\right|_{\tilde{f}_v^-}(f_v) \equiv (1-\lambda)\left[\tilde{f}_v \star \left(f_v \star \tilde{f}_v^- - f_d\right)\right] + \lambda\left(f_v - s_1\right) \equiv 0 \tag{8.5.4}$$

Similarly, given f_v, the minimum of J for $\tilde{f}_v^-$ is also obtained by setting the functional derivative to zero:

$$\left.\frac{\partial J}{\partial \tilde{f}_v^-}\right|_{f_v}\left(\tilde{f}_v^-\right) \equiv (1-\lambda)\left[f_v^- \star \left(f_v \star \tilde{f}_v^- - f_d\right)\right] + \lambda\left(\tilde{f}_v^- - \tilde{s}_1\right) \equiv 0 \tag{8.5.5}$$

Solving for equations 8.5.4 and 8.5.5 is however not straightforward. A naive approach could consist in solving it in as expressed above. For example, using transitivity and commutativity properties of the convolution, equation 8.5.4 could be rewritten as equation 8.5.6. Notice that the δ here accounts for the Dirac function, not for the location parameter of a stable distribution.

$$\left[(1-\lambda)\left(\tilde{f}_v \star \tilde{f}_v^-\right) + \lambda\delta\right] \star f_v \equiv (1-\lambda)\left(\tilde{f}_v \star f_d\right) + \lambda s_1 \tag{8.5.6}$$

Sampling f_v at regular intervals and using a finite difference scheme for the solver yields back a classical linear equation: $A \cdot X = B$. The main issue with the aforementioned method is its complexity. Writing K the number of sampling points for f_v, the inversion step for A yields a complexity of $\mathcal{O}\left(K^3\right)$. In addition, two convolutions have to be evaluated, but are in comparison very fast since they can be evaluated in $\mathcal{O}\left(K \log K\right)$. A secondary issue regards the conditioning of A which is hard to assess due to its dependency on $\tilde{f}_v$. Equation 8.5.5 can of course be interpreted similarly and leads to similar issues.

Due to these reasons, an alternative solver was preferred. Applying the inverse Fourier transform to equation 8.5.6 allows to replace convolutions by multiplications, thus yielding a straightforward solver. The most computationally expensive steps are now the Fourier transforms, hence implying a complexity of $\mathcal{O}\left(K \log K\right)$. The approach furthermore has the advantage of removing the requirement to estimate s_1 and $\tilde{s}_1$ from their characteristic functions [Čížek et al., 2005; Mittnik et al., 1999a]. The approach then simply summarises as alternatively solving equation 8.5.7 and 8.5.8:

$$\check{f}_v(\omega) = \frac{(1-\lambda)\left(\check{\tilde{f}}_v(\omega)\,\check{f}_d(\omega)\right) + \lambda\check{s}_1(\omega)}{(1-\lambda)\left(\check{\tilde{f}}_v(\omega)\,\widetilde{\check{\tilde{f}}_v^-}(\omega)\right) + \lambda} \tag{8.5.7}$$

$$\widetilde{\check{\tilde{f}}_v^-}(\omega) = \frac{(1-\lambda)\left(\widetilde{\check{\tilde{f}}_v^-}(\omega)\,\check{f}_d(\omega)\right) + \lambda\check{\tilde{s}}_1(\omega)}{(1-\lambda)\left(\widetilde{\check{\tilde{f}}_v^-}(\omega)\,\check{f}_v(\omega)\right) + \lambda} \tag{8.5.8}$$

In order to guarantee that the retrieved f_v and $\tilde{f}_v$ indeed represents probability density functions, one last issue should be tackled. The prior-information term, i.e. $\lambda\left(\|f_v - s_1\|_2^2 + \|\tilde{f}_v^- - \tilde{s}_1\|_2^2\right)$, do indeed ensures that both are close to a stable distribution, thus close to representing a probability density function. This is however not a strict requirement and the data-term might lead to negative values or non-normalisation for f_v and $\tilde{f}_v$. At each step, a re-normalisation procedure therefore takes place for f_v (respectively $\tilde{f}_v$) as stated in equation 8.5.9.

$$\begin{cases} \forall x / f_v(x) < 0,\ f_v(x) := 0 \\ f_v := \dfrac{f_v}{\int_{-\infty}^{\infty} f_v(t)dt} \end{cases} \tag{8.5.9}$$

8.5.3 Summary: Final Algorithm

According to the considerations given in previous sections, the final blind deconvolution method was defined as in algorithm 5. Although experience shows that the algorithm usually converges within a few iterations, it should however be stated that there is no formal proof of a convergence to a global optimum. Discussion on the convergence of such alternate minimisation schemes applied to blind deconvolution are however presented in [Chan and Wong, 2000] in the slightly less constrained case where the prior information term is based on smoothness (total variation) instead of a similarity argument.

The main conclusion was that the optimisation does empirically converge, although this depends on the initialisation. In this case, the initialisation is strictly defined by section 8.5.1 and the prior term involves a stricter constraint. These encouraging arguments are confirmed by practice: the algorithm converges within a few iterations.

Algorithm 5: Blind deconvolution algorithm.

Data: A convergence threshold: ϵ
Data: An estimate of f_d

Result: f_v and $\tilde{f}_v$

Initialisation:
Compute s_0 by means of the characteristic function method (section 8.4.4) Compute the pixel from the second flight-line corresponding to geo-location l: $\tilde{p}$
Initialise s_1 and $\tilde{s}_1$ from s_0 as in equation 8.5.3
$f_v := s_1$
$\tilde{f}_v^- := \tilde{s}_1$
begin
  while $\left\| \left.\frac{\partial J}{\partial f_v}\right|_{\tilde{f}_v^-}(f_v) \right\|_2 + \left\| \left.\frac{\partial J}{\partial \tilde{f}_v^-}\right|_{f_v}(\tilde{f}_v^-) \right\|_2 > \epsilon$ **do**
    Solve equation 8.5.7 for f_v
    Re-normalise f_v as in equation 8.5.9
    Solve equation 8.5.8 for $\tilde{f}_v^-$
    Re-normalise f_v^- as in equation 8.5.9
  end
end

8.6 Application

The concepts observed in the previous section were finally applied to the Kaufbeuren dataset, between flight-lines $4A$ and $4C$ which were the only two perfectly overlapping flight-lines. Illustrative results for band 10, 20 and 130 are displayed on figures 8.3 to 8.5. The convergence of these three examples was ranging from two to ten iterations, for a convergence threshold of $\epsilon = 10^{-3}$.

To start with, one should notice the quite small retrieved values for λ - respectively 0.030, 0.022 and 0.027. This indicates that the fitting to s_0 to f_d is indeed a valid choice as already seen in section 8.3.2. As a result, the convolution of f_v with $\tilde{f}_v^-$ indeed mimics f_d quite well.

A closer investigation of the results however reveals a slight artefact: $f_v \star \tilde{f}_v^-$ always slightly underestimates f_d near 0. This is in fact a result of the retained model. Stable distributions have a probability density functions defined on $\mathbb{R}$. On the other hand, f_d being a difference of reflectance, its probability density function is only ranging from -1 to 1. In other words, a stable distribution has a PDF which is more 'spread' than the actual f_d. Due to the fact that PDF are normalised, it is then logical that f_d is slightly greater than a stable distribution over $[-1, 1]$ and therefore greater than $f_v \star \tilde{f}_v^-$ which incorporates stable distributions into their constructions:

$$\int_{-1}^{1} f_d(t)dt = 1 = \int_{\infty}^{\infty} [f_v \star \tilde{f}_v^-](t)dt > \int_{-1}^{1} [f_v \star \tilde{f}_v^-](t)dt \tag{8.6.1}$$

A further point of interest regards the spread - and therefore the maximum value of f_v and $\tilde{f}_v$. When both differ (e.g. figures 8.3 and 8.4), f_v exhibits a greater maximum and a smaller spread than $\tilde{f}_v$. This can be explained by the already mentioned domain-definition issue in conjunction with the order of estimation used for f_v and $\tilde{f}_v$. Both functions are indeed initialised as stable distributions, hence with an infinite support. During the first iteration, f_v is updated first as a function of f_d which is compactly supported. It does therefore get a 'smaller' support size[1]. When $\tilde{f}_v$ is in turn estimated, f_v already accounts for the 'smaller' support and is therefore left with a wider spread, hence leading to the observed effect.

A workaround for both effects might be to work with censored LAS distribution (also called truncated LAS) [Mantegna and Stanley, 1994; Matacz, 2000]. The small amplitude of discrepancies however does not argue in favour of an increase of complexity.

8.7 Conclusions

This section has presented a methodology to estimate the uncertainty of the reflectance of each band. In practice, once f_v and $\tilde{f}_v$ have been retrieved, they can be directly used for two purposes.

The most straightforward application is to use them to compute the corresponding confidence intervals for the $L2$ reflectance, the difference of their median values[2] accounting for the relative biases.

A second use would be to use the corresponding distributions as input for any further Monte-Carlo error propagation. Since f_v is an estimation of the relative variability - i.e. the error distribution, it can be used to draw random reflectance samples before propagating them through any land-product model and collecting the output distribution. The latter would correspond to an uncertainty assessment as performed using a type-B Monte-Carlo method from the GUM - it

[1] The support size is actually infinite due to s_1: the density of f_v is however more grouped.
[2] Since they are close to stable distributions, their mean might be not defined.

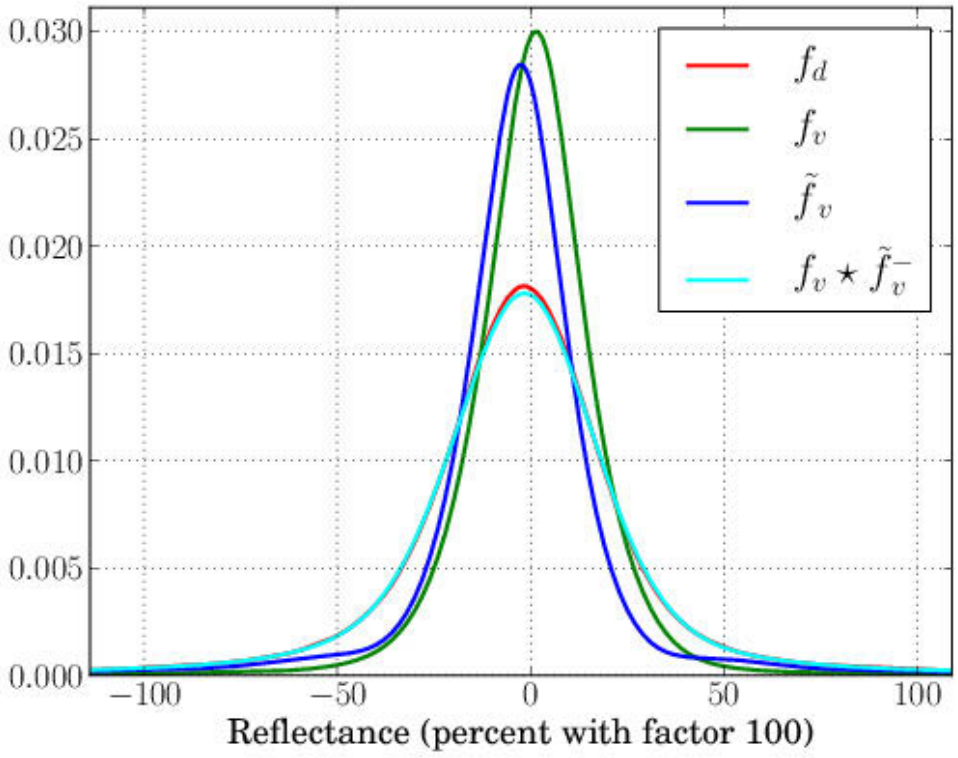

Figure 8.3: Retrieval of f_v and $\tilde{f}_v$ for the band 10 at 448 nm (Kaufbeuren 2012, flight-lines 4A and 4C).

would therefore only account for the variability introduced by the input data, not for the model induced one.

Another use could consist of using these estimates for an in-depth analysis of the variability itself. Having the distribution of the variability associated with the reflectance values, it is virtually possible to retrieve a rough estimate of the different contribution sources. This does however corresponds to a completely new issue - namely a blind-source separation problem [Mallat, 2009] - and is reviewed in the opening chapter of 10.

Finally, these estimates can be used to retrieve the effects produced by the pre-processing chain on the data. Since f_d is a global statistical representation, slight changes of input reflectance only slightly affects the matches retrieved by CPAM and should not affect f_d in other ways than by changing the actual reflectance. The retrieval of f_v and $\tilde{f}_v$ would then allow to study the pre-processing chain by means of a closed-loop transfer function.

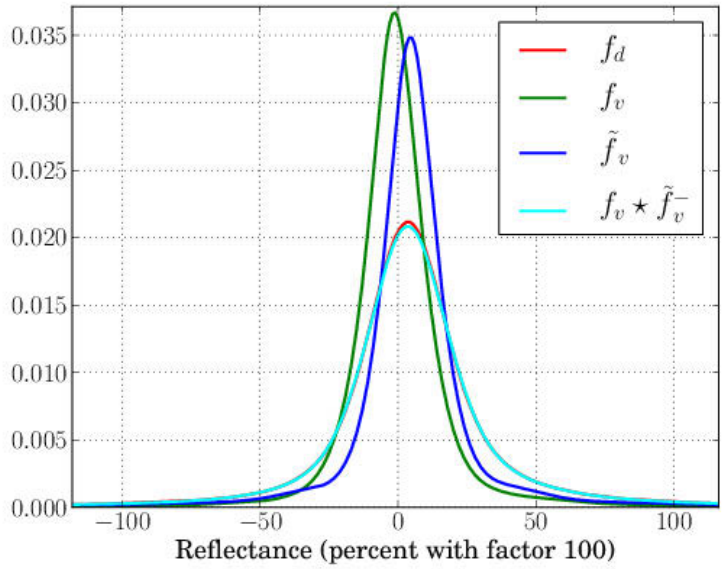

Figure 8.4: Retrieval of f_v and $\tilde{f}_v$ for the band 20 at 484 nm (Kaufbeuren 2012, flight-lines 4A and 4C).

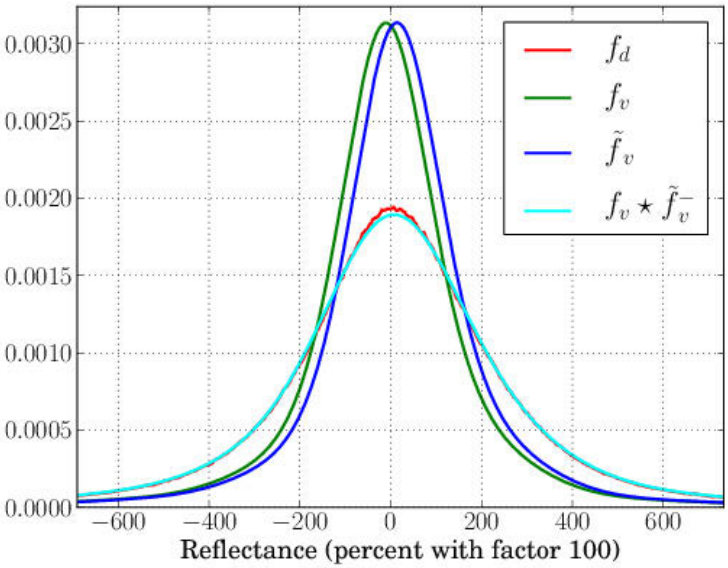

Figure 8.5: Retrieval of f_v and $\tilde{f}_v$ for the band 130 at 883 nm (Kaufbeuren 2012, flight-lines 4A and 4C).

Chapter 9

Summary

> The game of science is, in principle, without end. He who decides one day that scientific statements do not call for any further test, and that they can be regarded as finally verified, retires from the game.
>
> Karl Popper - Philosopher

Chapter Summary

This chapter critically reviews the main achievements of the work described in previous chapters in relationship to the issue statement.

9.1 Reminder: Problem Statement

Quality assessment should come along with any scientific data. In remote sensing this statement is not only driven by scientific integrity requirements, but also by several international initiatives such as *QA4EO* and - more specifically to hyperspectral data - by EUFAR. The German Aerospace Center, being an active member of both organisations, is naturally also striving towards this goal. Such quality assessment is usually understood to be twofold. The first aspect consists of providing data and products traceable to international standards. The second is related to providing quality indicators and confidence intervals with the data, thus improving knowledge about the reliability of data and products.

This works focuses on these issues in relation to the hyperspectral processing chain. Although the first issue has already been addressed by introducing suitable meta-data information and compliance with $ISO - 9001 : 2008$ guidelines, the second one is more open. The EUFAR initiative did propose several quality tags suitable for gaining a rough - theoretical - idea about the pre-processing quality. Such tags are, however, yet to be extended toward land-products. Furthermore, and even if they provide easy-to-interpret information, these remain quite limited, especially when compared with the recommendation for providing confidence intervals [QA4EO]. On the other hand, the complexity of the hyperspectral processing chain prevents any quality assessment based on errors propagation (e.g. Monte-Carlo error propagation). In addition, the involved pre-processing and correction models (atmospheric effects,...) are furthermore not completely error-free and need to be assessed too, thus preventing the investigations from being

solely conducted on the inputs' variability. Cross-validation methodologies allow for bypassing these limitations but are limited to the over optimistic case, where areas of interest are bright homogeneous lambertian surfaces. From this perspective, the goals of this work are threefold (see section 1.2):

- The EUFAR quality indicators have to be extended toward thematic land-products based on remote sensing data.
- The processing chain should be improved to limit the corresponding uncertainties associated with mapping results.
- Data from $L2$-level and higher should be provided with confidence intervals.

Besides these requirements, operational issues should be considered too. To start with, the corresponding quality assessment has to be included within the DLR hyperspectral processing chain. It therefore has to be both generic enough to be run on any data-set as well as automated enough to limit human interactions to a minimum.

9.2 Review of Proposed Solutions

From the needs and constraints stated in the previous section, it appears that the common uncertainty propagation methods cannot be used as such, should it only be due to the complexity of the processing chain. On the other hand, due to the design of hyperspectral sensor systems, data is usually acquired as several partially overlapping flight-lines, hence producing information redundancy. This redundancy and the underlying discrepancies can in turn be used to assess the measurement's repeatability.

Such an assessment is, by essence, specific of the investigated scene and therefore tuned to it. Besides, the retrieved discrepancies are relative ones (i.e errors), hence complementing existing approaches retrieving biases instead of errors. This approach can finally be performed on virtually any scene, thus providing a generic tool for error assessment.

Although appealing, this approach is linked to a data limitation which had to be raised. Two consecutive flight-lines are in practice never completely co-registered, hence preventing a direct comparison of their results. As a first step, a semi-automatic co-registration algorithm was therefore proposed to overcome this limitation. It retrieves pairs of pixels corresponding to the same ground footprint based on local spectral similarities. To start with, for each pixel from the first flight-line, the most similar one from the second flight-line is searched for. Once it has been obtained, the inverse procedure takes place: the first flight-line is searched for its most similar correspondence. If the start pixel is indeed retrieved, the pair is kept as being two observations of the same footprint. On the other hand, if the second step does not lead to the start pixel, the pair is considered as unsafe and discarded. Finally, the spectral similarity is not computed on a pixel to pixel basis, but on a local basis, thus limiting the impact of the noise. As a result, retrieved pairs of pixels are mostly valid. The validation step indicated that, in the worst case, less than 10% of the retrieved pairs are doubtful. In most cases, it is less than 2%.

This quite harsh rejection of untrusted pairs has several advantages. Over large homogeneous surfaces, where the lack of contrast leads to a higher sensitivity to noise, the proposed algorithm is likely to retrieve fewer matches. As a result, matches are denser over heterogeneous areas, hence providing an estimation of the variability in the worst case. In particular, neighbouring effects and differences in the pixels footprints have their effects maximised. Since other empirical quality assessment techniques are usually focusing on homogeneous areas, this property helps to complement them.

On the other hand, it might occur that an area of interest is left aside: this is, that too few matches could be retrieved to allow any further analysis. In order to tackle this potential drawback, a thin-plate-splines interpolation methodology has been proposed. The presented approach consists of an interpolation based on physical considerations while maintaining a reasonable algorithmic complexity.

Once pairs of corresponding pixels have been retrieved, it is proposed to use them for performing a data-driven quality assessment. It is in particular possible to assess any land-product computed on a pixel-wise basis. This is achieved by comparing results obtained for each pair of pixels, hence translating the quality assessment issue into an inter-rater reliability assessment. In order to tackle the wide range of possible land-product types (classification, quantitative analysis) it is proposed to use the Krippendorff's α reliability coefficient, thus extending the EUFAR quality indicators up to land-products. Furthermore, and since in practice large amount of pixels pairs are retrieved, it becomes possible to provide confidence intervals of these criteria based on bootstrapping. Besides being more stable with respect to data, this approach is generic in terms of both the type of data which can be handled (nominal, absolute, ...) and in terms of its flexibility with respect to the number of observations which can be compared. Using this approach, EUFAR quality tags are therefore not only extended towards land-products, but also towards recommendations from QA4EO. Last but not least, these results are based and therefore tailored on the data itself.

In a second step these reliability coefficients can be used to improve land-products generation. The mapping applications indeed depend on one or more parameters. Maximising the retrieved reliability over the set of admissible parameters therefore leads to an improvement of the land-product. For the presented exemplary data-set, this yields a reliability improvement of up to 24%.

At this stage, one final gap remains: there is no confidence interval available for the reflectance data itself. In order to address this, a blind-deconvolution framework was proposed to retrieve individual uncertainties from the discrepancies between two perfectly overlapping flight-lines. The ill-posedness of the problem is raised based on both theoretical and empirical considerations, leading to the modelling of uncertainties as being similar to the levy alpha stable distribution. As a result, an estimator of the errors' probability density function is obtained.

9.3 Critical Review of the Achievements

Section 9.2 has reviewed how the proposed approach solves the issues faced when assessing the quality of hyperspectral systems, thus allowing a complete quality assessment including thematic products. However, this approach should not be considered as a strictly standalone solution. The next paragraphs review its underlying hypotheses and limitations.

This work analysed the data and product quality in terms of reproducibility. As a result, no statement over the systematic biases is made: focus is set on relative errors in the worst case scenario. The approach therefore fills a gap in quality assessment methodologies but does not provide a complete framework on its own: it would therefore be used in conjunction with other techniques such as cross-calibration (or vicarious calibration) over homogeneous areas. A second element of interest is the actual uncertainty retrieved. Except for the estimation of the reflectance, the comparison of two flight-lines does not only account for the system uncertainties: it also accounts for the scene variability (e.g. with respect to viewpoint changes). This is a valuable property when trying to state whether a specific parameter could be mapped with a specific accuracy. On the other hand, it also 'hides' the uncertainty of the system into this more global term. Depending on the final users needs, this might be both an advantage or a

disadvantage.

A further point regards the improvement of mapping results based on a data-driven optimisation of the parameters for land-product generators. Even though the results have improved, the reflectance data is not - and cannot - be made more reliable. This remains an open point which may be addressed in further work, for example by making use of blind source separation algorithms.

The blind-deconvolution approach for reflectance uncertainty estimation induces slightly different limitations. To start with, the main contributor to scene changes - the change of viewpoint - has been discarded. Even if some scene-induced uncertainty is involved in the computations, they are negligible and can moreover be linked to the sensor's system limitations. For example, the slight changes of pixels' footprints from flight-line to flight-line can be seen as a limitation due to the sampling performed by the sensor. The retrieved information is therefore slightly more precise, the drawback being a stronger requirement in terms of data availability: namely two completely overlapping flight-lines.

Beside these points, the proposed approached can be used in a quite automated fashion on virtually any data-set. Since it is data-driven and therefore independent from the details of the processing chain, it can be easily incorporated in the existing processing chain and is unlikely to be affected by further changes or improvement. Finally, it does not require any additional data[1] and can therefore be used for reprocessing archived data, thus ensuring a backward compatibility. These properties ensure a complete compliance with the aforementioned operational constraints.

9.4 Outlook

As final words, it should be underlined that a complete quality assessment of the hyperspectral system is by no mean fully achieved. The following points are treated in more details in Chapter 10.

To start with, it should be stressed that results presented in chapter 8 are only computed on a band-wise basis: the reflectance's variability is therefore not addressed over the full spectrum. In particular, relationships amongst the variability of different bands can not be retrieved. In order to account for such effects, higher-level frameworks should be brought in: section 10.2 reviews this issue in more details.

A second step would consist in extending the analysis of the variability: getting an estimate of the probability density function of the reflectance's variability is a useful outcome. On the other hand, once it has been made available, the next step would consist in retrieving the uncertainty sources and in particular which ones are contributing the most. Getting such information would indeed allow, on the longer run, to identify the bottle necks of the system, hence leading to potential improvements. This is further discussed in section 10.3.

Besides these long-term scientific tasks, a further short-term technical task should also be made. This work has indeed developed a complete range of theoretical methodologies. These have, in a second stage, already been implemented and tested on several data-sets. The next step would then consist of inserting them into the DLR's processing chain in order to systematically test them on data-sets from upcoming flight-campaigns.

[1] Except for the reflectance uncertainty estimation.

Chapter 10

Possible Extensions

> I have discovered a truly marvellous proof for this, but this margin is too small to contain it.
>
> Pierre de Fermat - Scientist

Chapter Summary

This chapter proposes several potential extensions of the work presented in the previous sections. To start with, considerations about the representation of spectral variability are discussed. In a second stage, possible extensions of a deeper analysis of the uncertainties sources are reviewed.

10.1 Introduction

Although several problematic issues have been addressed within this work, the topic of quality assessment for hyperspectral systems is far from been closed, even when considering parallel works. This chapter reviews some possible future research directions addressing some of the remaining aspects.

Section 10.2 presents some considerations on the statistical representation of the spectra and their variability. Section 10.3 then reviews some possibilities for a finer decomposition of the uncertainty sources. Finally, section 10.4 presents concluding remarks on this work.

10.2 Better Description of a Spectral-Shape Variability

Readers may have noticed that most of the analysis of the spectral content is performed on a 'per band' basis, thus discarding any information contained in the spectral shape. To take a single example, the spectral angle mapper compares two spectra in terms of a scalar product - the arc-cosines being just a diffeomorphism providing a non-linear scaling. This scalar product removes a quite large portion of the information contained in the spectra: namely all potential correlations between neighbouring bands.

From a mathematical point of view, this can be seen as an issue from the spectral representation itself. As seen in section 2.4.1, spectra are usually considered as being vectors from $\mathbb{R}^B$, B being the number of sampled bands. However, only a tiny subset of this B dimensional space is used. Since it physically corresponds to reflectance, the actual space is $[0;1]^B$. Furthermore, signal from adjacent channels are correlated in practice: un-correlated spectral bands

are actually considered as being faulty [Bachmann et al., 2010a, 2011a]. This yields - among other considerations - that two consecutive bands should not be considered as being orthogonal in terms of information content. This is however the implication done when considering spectra as elements from $\mathbb{R}^B$.

Furthermore, the variability of a spectrum is known to not be strictly band dependent. For example, vegetation exhibits shifts of its features from band to band when subjected to stress (see: figure 10.1). Performing an analysis of the full spectral variability would therefore make much more sense than just analysing each band independently. In terms of pure noise, and not physically induced variability, such phenomena do not seem to have been well studied. Some authors do however hint that there is at least a slight correlation between the noise from band to band [Aiazzi et al., 2006, 2010; Farzam and Beheshti, 2011].

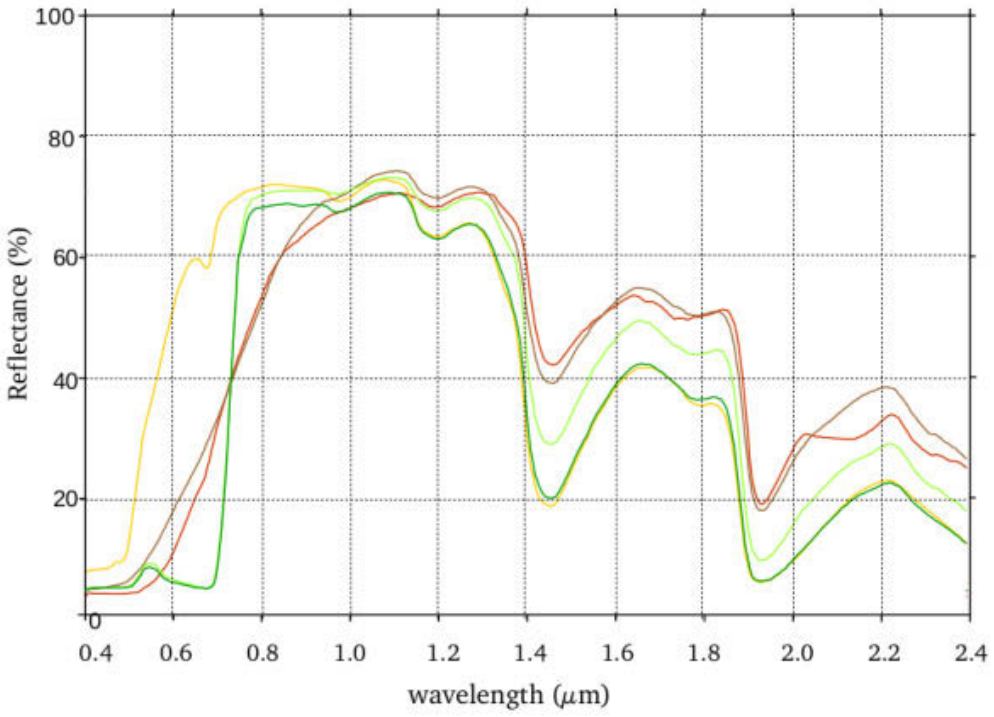

Figure 10.1: Spectral variability of vegetation when submitted to stress (yellow) or drought (brown) as opposed to typical spectra of healthy vegetation (green).

From a practical point of view this lack of modelling has an actual impact. The 'shape' of each spectrum - defined by the relationship between different spectral channels - is valuable when retrieving quantitative information about the surface materials (e.g. [Bayer et al., 2012; Ehrler, 2014]). The current modelling of spectra as vectors from $\mathbb{R}^B$ is therefore too limited and alternatives have to be retrieved. The following sections reviews some alternative views that could be used for further analysis.

10.2.1 Spectra as Vectors: Principal Component Analysis

The principal component analysis - or PCA - is also called single-value-decomposition or Karhunen-Loéve transform. It aims at translating a vector set from $\mathbb{R}^B$ into a new basis where the corresponding components are not linearly correlated. Given a set of vectors from $\mathbb{R}^B$, a PCA is simply performing a single value decomposition of the associated covariance matrix. The associated eigenvectors then form a new basis where the vectors can be decomposed with non linearly correlated components. In general, the components associated with very low eigenvalues are considered to represent measurement noise and are therefore discarded, the remaining eigenvalues can be seen as representing the 'amount of standard deviation' associated with the direction of their related eigenvector. In hyperspectral remote sensing, PCA is a commonly used tool, for example when retrieving end-members of a data-set [Rogge et al., 2007].

A straightforward alternative to the 'per-band' analysis presented in section 8 could be to perform a classification for a set of 'broad' classes (bare soil, concrete, vegetation,...). A PCA could then be performed on the discrepancy spectra - i.e. f_d computed over the whole spectrum instead of over a single band. A deconvolution process could then take place on the coefficients from the basis retrieved by PCA, thus taking into account linear correlations within the data.

Although straightforward to compute and analyse, PCA has several drawbacks. As already stated, it only accounts for linear correlations. The correlations are furthermore purely based on the data: it is unlikely that the new basis computed by PCA over a vegetated area would correspond to the one computed over a concrete area, hence the need to perform a pre-classification. Such a work-around would in turn involve further issues such as the choice of the rightful classes, how to deal with mixed pixels, etc. As a result, this method is unlikely to be really useful.

10.2.2 Spectra as Functions: Wavelet Analysis

An alternative to PCA would be to consider each spectra as what it actually is: a function of the wavelength whose value is sampled over each band. It should be noted that such a viewpoint is not uncommon. For example, the well known Savitzky-Golay filtering commonly applied to field spectra [Bayer et al., 2012; Ehrler, 2014] actually considers the input spectra as a function sampled on a regular grid before fitting it to local polynomials [Savitzky and Golay, 1964].

In general, the whole point in considering spectra as functions consists of taking advantage of an ad hoc orthonormal functional basis. In this case, the following only focuses on the principles of such an analysis without stepping into the mathematical details as in [Allaire, 2005; Mallat, 2009; Tarantola, 2005]. A convenient basis for such an analysis is the wavelets analysis: most of the following originate from [Mallat, 2009]. In the following, each spectra are defined as a function from $L^2(\mathbb{R})$ and are written $f : \lambda \rightarrow f(\lambda)$.

Principles

The basic idea behind the use of a wavelet basis consists of decomposing each spectrum as the sum of an approximation element taken in a subset $V \subset \mathbb{L}^2(\mathbb{R})$ and increasing levels of details taken from elements from $\mathbb{L}^2(\mathbb{R})$: $\{W_k\}$, each of these subsets being mutually orthogonal.

Given a basis of V: $\{\phi_n\}$ and basis for W_h: $\{\psi_{k,n}\}$, f can be decomposed by means of the classical scalar production on $\mathbb{L}^2(\mathbb{R})$:

$$f(\lambda) = \sum_n \langle f, \phi_n \rangle \phi_n(\lambda) + \sum_{k=0}^{\infty} \sum_n \langle f, \psi_{k,n} \rangle \psi_{k,n}(\lambda) \tag{10.2.1}$$

The ϕ_n is here an integer shift of a compactly supported function ϕ, also called scaling function. It can be shown that such a decomposition actually exists and that, for each (k, n), a

valid choice for $\psi_{k,n}(x)$ is $1/\sqrt{2^k}\psi\left(\left(x-2^k n\right)/2^k\right)$, where ψ is the so-called mother wavelet and is linked to ϕ which is in turn further constrained by some other considerations [Mallat, 2009].

Neglecting details that are too small to be actually retrieved by the spectral measurement, and up to a rescaling to 'optimise' the approximation representation under the assumption that f is also compactly supported, equation 10.2.1 can be simplified into equation 10.2.2:

$$\begin{aligned} f(\lambda) &= \langle f,\phi\rangle\phi(\lambda) + \sum_{k=0}^{K}\sum_{n}\langle f,\psi_{k,n}\rangle\psi_{k,n}(\lambda) \\ &= a\phi(\lambda) + \sum_{k=0}^{K}\sum_{n} d_{k,n}\psi_{k,n}(\lambda) \end{aligned} \tag{10.2.2}$$

In general, f is however only known on a finite number of irregular sampling points, thus practically preventing the direct estimation of the aforementioned scalar products. If the sampling had been performed on a regular grid, some fast algorithms could be used to estimate the coefficients $d_{k,n}$ and a: the so-called fast wavelet transform (also called discrete wavelet transform of DWT).

Wavelet interpolation

In spectroscopy, the use of DWT is however prevented due to the irregularity of the sampling bands: the spectral channels are typically not equally spaced. SWIR bands are typically wider and further apart than the VIS ones. Furthermore, bands located in the atmospheric absorption domain are discarded once the atmospheric correction step has been performed. This implies that efficient computation methods such as the discrete wavelet transform can not be used.

In [Bernard et al., 1998; Bernard, 1999], an alternative to this limitation was proposed by iteratively selecting the wavelets $\psi_{k,n}$ close enough to a sampling point to get their coefficient retrieved. The more compact a set of sampling points is, the more detailed information can be retrieved. The spectral bands are iteratively parsed, and increasing levels of details associated with their locations are iteratively added. Given B bands, located at wavelengths $\{\lambda_b\}$, a total of $B-1$ locations and levels of detail can be retrieved: $\{(k_b,n_b)\}$. Since the values of f are known at each λ_b, this yields a system of B linear equations with B unknowns - a and $\{d_{k_b,n_b}\}$ - which can then be solved through a linear system inversion.

$$\left\{\begin{array}{rcl} f(\lambda_1) &=& a\phi(\lambda_1) + d_{k_1,n_1}\psi_{k_1,n_1}(\lambda_1) + \cdots + d_{k_B,n_B}\psi_{k_B,n_B}(\lambda_1) \\ \ldots &=& \ldots \\ f(\lambda_b) &=& a\phi(\lambda_b) + d_{k_1,n_1}\psi_{k_1,n_1}(\lambda_b) + \cdots + d_{k_B,n_B}\psi_{k_B,n_B}(\lambda_b) \\ \ldots &=& \ldots \\ f(\lambda_B) &=& a\phi(\lambda_B) + d_{k_1,n_1}\psi_{k_1,n_1}(\lambda_B) + \cdots + d_{k_B,n_B}\psi_{k_B,n_B}(\lambda_B) \end{array}\right. \tag{10.2.3}$$

As a concluding remark, it should be noted that given the coefficients, an estimate of f is then available over the complete set of admissible wavelengths, not only on the bands it was sampled on. Considerations about the corresponding error are discussed in several different cases in [Bernard, 1999; Daubechies et al., 1999; Deslauriers and Dubuc, 1989; Feichtinger and Gröchenig, 1994].

Wavelet choice

Once this projection setup has been defined, a final step would be to chose carefully which approximation and mother wavelet should be used for this task. As stressed out in [Mallat, 2009], two main properties should be considered: the wavelets' support size and their number of vanishing moments[1] which are antithetic.

The support size is in our case a determining factor: spectra are likely to exhibit singularities: in order to avoid having them pollute a large amount of neighbouring wavelets, wavelets should have an as small as possible support. On the other hand, the number of vanishing moments is 'only' indicative of how well summarised the information is. In this study, this is actually only a secondary concern.

In [Daubechies, 1988], a family of wavelets of minimum support size given a number of vanishing moments has been proposed. Their main issue lies in the fact that they do not have an analytical expression[2] except for a number of vanishing moments equal to zero (also called Haar wavelet, see equations 10.2.4). Luckily, this also corresponds to the most compactly supported orthogonal wavelet. On the other hand, it has no vanishing moments, hence it does not perform well when approximating smooth spectra.

$$\phi : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R} \\ \lambda & \rightarrow & \begin{cases} 1 & \text{if} \quad \lambda \in [0;1] \\ 0 & \text{else} \end{cases} \end{cases} \qquad \psi : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R} \\ \lambda & \rightarrow & \begin{cases} -1 & \text{if} \quad \lambda \in [0;1/2[\\ 1 & \text{if} \quad \lambda \in [1/2;1[\\ 0 & \text{else} \end{cases} \end{cases} \tag{10.2.4}$$

Practical Use

By combining equations 10.2.3 and 10.2.4, it is possible to retrieve coefficients for every single spectrum. This would in turn allow several new investigations, as in chapter 8, but working on wavelet coefficients instead of bands values. Working on such wavelet coefficients however allows a quite different viewpoint on the issue.

To start with, the wavelet coefficients are modelling not only the bands themselves but also their relationships with a given neighbourhood. In other words, they are not only representing their values, but also the underlying frequencies. The last compromise is furthermore done in the limit case of the uncertainty principle: this includes the smallest frequency uncertainty given a location uncertainty [Mallat, 2009]. These properties in turn allow for studying uncertainties in terms of 'band-to-band' interactions. This would be especially interesting in terms of their effects on band ratios.

A slightly more practical implication would be to allow the comparison of sensors sharing similar designs but not necessarily the same exact spectral bands. Provided some of the retrieved wavelets are the same for both sensors, the variability of their coefficients retrieved for both sensors could theoretically be compared, thus opening the way to an objective comparison of sensors.

Finally, in terms of confidence intervals representation, as recommended by [QA4EO, 2010], presenting the variability in terms of wavelet coefficients permits to depict the variability of the spectra as a function, thus making it usable in practice. For example, as illustrated in table 10.1, investigating only the variability from band to band would lead to a confidence interval for both

[1] Having p vanishing moments means that the wavelet is orthogonal to any polynomial of degree $p-1$.

[2] It is however possible to use them easily in case of a DWT. It should also be noted that more complex refinement of the methodology allow for more flexibility: so-called multiwavelets [Geronimo et al., 1994; Strang and Strela, 1995].

bands of about ± 7, thus implicitly stating that the discrepancy of 2 between each pair is lost in the variability. However, when working on wavelet coefficients, it can be directly retrieved that there is only a global amplitude variability associated with the projection on ϕ, thus allowing us to retrieve the actual usability of the data.

Run	$f(\lambda_1)$	$f(\lambda_2)$	a	$d_{0,0}$
A	1	3	2	1
B	8	10	9	1
C	15	17	16	1
CI	± 7	± 7	± 7	± 0

Table 10.1: Illustration of the advantages of representing the variability between the runs A, B and C in terms of wavelet coefficients when dealing with confidence intervals.

In terms of drawbacks, it should be pointed out that the additive model as stated in 8.2, cannot directly be used any more. For example, the approximation coefficient a might be modelled as being affected by an additive variability. On the other hand, the most detailed coefficients' variability should be modelled by means of higher-level models since they might represent intra-spectra variability (i.e. when comparing two spectra f and $\tilde{f}$) as well as the spectrum intra-variability (i.e. the amount of noise occurring from band to band). Although not impossible, this is likely to yield a high load of work, when it comes to analysing and modelling the corresponding uncertainties.

10.2.3 Spectra as Shapes: Q-Functions

A higher level research direction might be to directly depict spectra as shapes - such as for example the $\mathcal{Q}$-function framework. Although this implies an even more complex background and computations than the wavelet-based analysis, the main advantage would be to allow a full shape variability description that would likely be more adequate to describe uncertainties, as well as effects displayed on figure 10.1.

This section quickly reviews the main concepts involved in the $\mathcal{Q}$-functions framework (more detailed information, especially about the mathematical concepts can be found in [Joshi et al., 2007; Klassen et al., 2004; Mani et al., 2010; Srivastava et al., 2005, 2011]). The following paragraphs are also referring to several mathematical concepts whose exact definitions can be found in mathematics text books. For readers lacking this understanding, 'with the hands' definitions are given below, the exact definitions requiring the introduction of advanced concepts and notions. The relevant explanations are:

- Hilbert space: this is a generalisation of the classical three-dimensional Euclidean space to spaces having a larger or even infinite dimensionality.
- Manifold: it can be seen as a generalisation of the notion of surface for larger or infinite dimensions, where local properties can be similar to ones encountered on a Euclidean space.
- Tangent space of a Manifold: this can again be seen as the generalisation of a tangent.
- Riemanian metric (on a differentiable manifold): this is a bi-linear symmetric mapping defined on the the tangent space. A Riemanian manifold is a manifold for which a Riemanian metric could be defined.

- Geodesic path between two points of a Riemanian manifold: this is the generalisation of the notion of 'straight line' to the non flat surface defined by the manifold.

To start with, the spectra are now considered as a bi-dimensional parametric curve as stated in equation 10.2.5. One could note that - even if this is of little use in terms of spectrometry - this representation is more generic than the pure functional description from section 10.2.2.

$$\gamma \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R}^2 \\ t & \rightarrow & \gamma(t) = \begin{pmatrix} f(t) \\ \lambda(t) \end{pmatrix} \end{cases} \tag{10.2.5}$$

In its basic form, $\mathcal{Q}$-functions are based on the square root velocity function of the curve, where $\| \bullet \|$ is the Euclidean distance: see equation 10.2.6.

$$\bar{q}(t) = \begin{cases} \frac{\partial \gamma}{\partial t}(t) \Big/ \sqrt{\left\| \frac{\partial \gamma}{\partial t}(t) \right\|} & \text{if} \quad \left\| \frac{\partial \gamma}{\partial t}(t) \right\| \neq 0 \\ 0 & \text{else} \end{cases} \tag{10.2.6}$$

Beside several interesting properties, such functions are well defined and continuous, even on singular points, and are iso-metric to the original curve, given the classical scalar product in L^2. This function is finally rescaled to obtain the 'final' $\mathcal{Q}$-function: $q = \bar{q} / \int \|\gamma(t)\| dt$. It can now be proved that the corresponding curves are elements of a hypersphere from the Hilbert manifold of bi-dimensional parametric functions which is a Hilbert sub-manifold: it is called pre-shape space and noted $\mathcal{C}^0$. The final step consists of removing the influence of re-parametrisation and rotation by forming the quotient space of $\mathcal{C}^0$ modulo re-parametrization and rotations. The resulting space being called shape space or $\mathcal{S}$. It is finally possible to show that $\mathcal{S}$ is a Riemanian manifold, hence representing a 'shape space'.

Once this shape space has been defined, and given a set of spectra, one can obtain their correspondence in $\mathcal{S}$, thus compare them in terms of their geodesic path. It is in particular possible to compute the Karcher mean of a set of shapes $\{q_k\}$ as being the manifold point minimising the sum of its distances[3] to each q_k. Given this 'mean-shape' - μ - one can then project each q_k to the tangent space to the manifold taken at μ. As it is linear, usual statistics can then be performed there, before being re-projected back.

As a conclusion, this $\mathcal{Q}$-function framework allows for the definition of statistic and distances on shapes. Applying it on the shapes defined by spectra would lead the way to reliably perform statistics susceptible to model the spectral variability. It should in particular be able to model shifts in both depth and wavelength of the spectral features. Although this could have been partially done using the wavelets basis, $\mathcal{Q}$-functions allow for a much more precise modelling. One example could consist of its ability to smoothly merge two absorption features into one, where the wavelet analysis would model it as two small details disappearance to the profit of the appearance of one larger detail (see figure 10.2). On the other hand, such a model is much more computationally intensive.

Both proposed frameworks are however not completely disjoint. Statistics in the tangent space also require a suitable basis in the corresponding functional space. This is usually done numerically, but one could think about introducing the use of a wavelet basis instead.

As stated, the methodology is not directly adapted to spectroscopy. For example, the invariance to rotation is discussable when applied to the spectral analysis. Two linear spectra, one oblique and one horizontal, have different meanings. All the same, the implicit translation

[3] In terms of the corresponding Riemanian metric.

Figure 10.2: Deformation between three pairs of shapes according to the geodesic paths between them. Bumps are only created where needed, existing ones being shifted and expanded.

invariance due to the derivation in equation 10.2.6 might be discussable when comparing a linear spectrum with low reflectance (e.g. coal) to a spectrum with linear high reflectance (e.g. spectralon ®).

A final remark should be made about the spectral angle. When spectra are normalised, thus being represented as elements in hypersphere in $\mathbb{R}^B$, the spectral angle, expressed in radians, is nothing less than a geodesic distance between two spectra on this sphere. This property is likely the explanation of its success, although there does not seem to have been any work done towards taking advantage of the underlying tangent planes.

10.2.4 Conclusions

This first set of extensions aim at providing a more suitable framework for the analysis of spectral variability. It is indeed believed that a necessary condition for performing a reliable analysis of the hyperspectral data consists of choosing a data-representation framework in compliance with the meaning of the data itself. Should it only be for the sake of presenting the variability of the data in a usable fashion instead of performing a quality assessment for the sake of quality assessment (see table 10.1).

The proposed methodologies have been presented from a theoretical point of view. They are however linked to a quite heavy background and would require a large amount of work before being fully operational. The wavelet analysis would, in general, be the more straightforward in both theoretical and computational complexity. The $\mathcal{Q}$-function framework, on the other hand, is more detailed but also much more complex.

A final practical aspect lies in the community background. The modelling of spectra as elements from $\mathbb{R}^B$ is widely spread and several specific methods have been developed to address the underlying issues when applied to specific applications (classifications, end-member extractions,...). It seems unlikely that this whole know-how could be sacrificed for the sake of an ad-hoc quality assessment. Just as for the use of a 'hidden' notion of geodesic path in the spectral-angle, it might however happen that these methodologies could be incorporated within a wider scope.

10.3 Blind Source Separation

Independently from the chosen spectral representation, results from chapter 4 could lead to another direction of analysis. The basic idea consists of representing the measured uncertainty as the sum of contributions from different influences. Under a linear additively assumption[4], this yields equation 10.3.1 for each pair of pixels retrieved by CPAM. Here f is the observed variability, f_s is the contribution of the source s, S being the total number of sources. Finally η is an estimation noise, while u_s is the mixing parameter associated with source s .

$$f = v - \tilde{v} = \sum_{s=1}^{S} u_s f_s + \eta \tag{10.3.1}$$

In general, this is an extremely ill-posed problem, both $\{f_s\}$ and $\{u_s\}$ as well as η being unknown. Under such conditions, it is very unlikely that all different sources could be identified. In fact, only independent sources or at least sources having approximation supports which are not overlapping can be retrieved [Mallat, 2009]. In practice there are luckily several broad groups of influences which might be partially discriminated (see figure 10.3):

- BRDF effects, inducing multiplicative factors
- Pure noise affecting all bands, independently from the observed signal
- Differences in pixels' footprints between two flight-lines

Under the assumption of statistical independence of the sources, several methods have been elaborated based on the so-called independent components analysis. It aims at decomposing the signal as the sum of the most independent sources [Cardoso, 1999; Nguyen and Zheng, 2011; Zarzoso et al., 2010]. This independence condition, as already discussed in chapter 8, section 8.2.3, is however not always met.

Alternatives might reside in attempting to use some prior modelling to build a dictionary where each source would have a 'sparse' representation [Bofill and Zibulevsky, 2001; Mallat, 2009; Zibulevsky and Pearlmutter, 2001]. Several methods have already been developed in the fields of sound analysis [Yilmaz and Rickard, 2004] and wireless communications [Jourjine et al., 2000]. These approaches are linked with decompositions on wavelet basis, and are linked to some un-mixing approaches in hyperspectral remote sensing [Bioucas-Dias et al., 2012].

Intensive research however remains to be done as to the feasibility of theses results with respect to practical uses.

10.4 Conclusions

The two previous sections have pointed out that the current status is still quite far from reaching a complete quality assessment, even though several steps have been made toward this.

In particular, the issue of the description of a spectrum - as well as the associated uncertainty - is assumed to be one of the major bottle necks. Further work should therefore be done in order to develop a suitable description, thus allowing an in-depth analysis of the uncertainties. From the three proposed alternatives, the raw PCA analysis seems to be comparatively crude and limited. The wavelet-based framework would be comparably fast and reliable alternative. On the other-hand, a dedicated shape analysis framework is likely to produce better results, at the cost of a much higher complexity.

[4]Note the analogy with the general central limit theorem, as expressed in section 8.3.2.

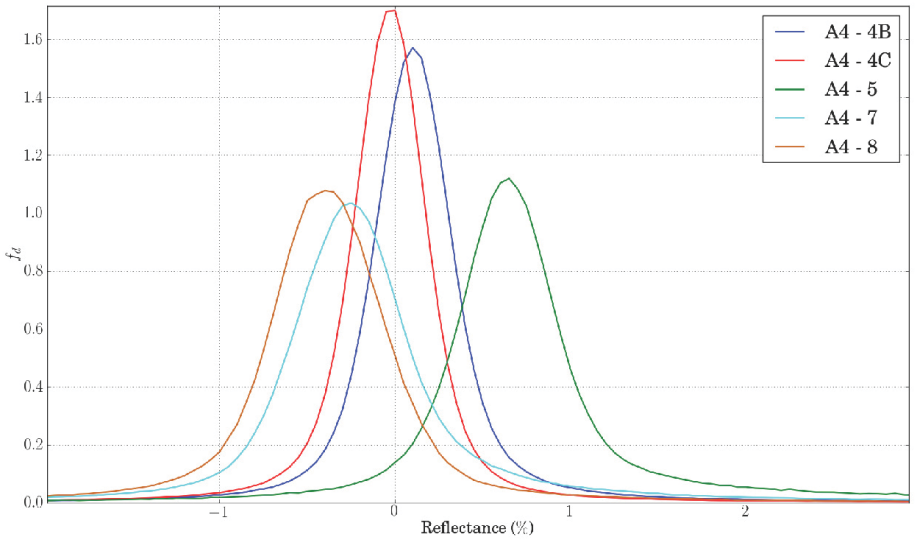

Figure 10.3: Illustration of BRDF effects with respect to variability on the Kaufbeuren 2012 dataset (band 10). Variability induced by differing viewing angles ($4A-5$ or $4A-7$ & $4A-8$) is different from the variability obtained when the same viewpoints are used ($4A-4B$ & $4A-4C$) and should therefore be discriminative.

Another research direction would be to aim towards a more detailed description of the uncertainty sources. Although it is very unlikely that a solution leading to a detailed description could be achieved, it might be possible to retrieve a rough estimator. Such information could in turn be used to allow a more targeted processing chain improvement.

Finally it might be interesting to get a finer understanding of the relationship between sensor and scene induced variability. For example [Reulke and Weichelt, 2012] have shown that the noise equivalent radiance actually depends on the radiance. A better understanding of such a link is likely to open new doors.

Beside these technical considerations, it should be noted that the proposed methods are not limited to multi- and hyperspectral data. It is indeed virtually possible to extend it towards mono-band data (e.g. gray-scale imagery, elevation models). As presented in chapter 4 several methods exist to search for correspondence on single-band imagery. The remaining chapters stay unchanged when limiting the number of bands to one. In particular, chapters 6 and 7 are valid for any type of rasterised geographical-data, provided an overlap indeed exits. Finally, results presented in chapter 8 could be used for modelling any additive variability - eventually at the cost of altering the 'a priori similarity to stable distribution' in order to take into account different types of variability.

Finally, as already stated in section 8.7, outputs from chapter 8 could be used as a feed-back loop for analysing impacts of changes of the pre-processing chain in terms of reflectance.

Bibliography

A. Abdel-Hakim and A. Farag. Csift: A sift descriptor with color invariant characteristics. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 1978–1983, 2006. doi: 10.1109/CVPR.2006.95.

P. Abel, B. Guenther, R. N. Galimore, and J. W. Cooper. Calibration results for NOAA-11 AVHRR channels 1 and 2 from congruent path aircraft observations. *Journal of Atmospheric and Oceanic Technology*, 10(4):493–508, August 1993. ISSN 0739-0572. doi: 10.1175/1520-0426(1993)010.

D. Adam. Gravity measurement: Amazing grace. *Nature*, 416(6876):10–11, March 2002. doi: 10.1038/416010a.

E. H. Adelson. Perceptual organization and the judgment of brightness. *Science*, 262(5142): 2042–2044, Decembre 1993. doi: 10.1126/science.8266102.

AFS, Aviation Flight Standards Service, U.S. Department of Transportation, and Federal Aviation Administration. Advisory circular 120-82: Flight operational quality assurance. Advisory Circular 120-82, Federal Aviation Agency (FAA), December 2004.

B. Aiazzi, L. Alparone, A. Barducci, S. Baronti, P. Marcoionni, I. Pippi, and M. Selva. Noise modelling and estimation of hyperspectral data from airborne imaging spectrometers. *Annals of Geophysics*, 49(1), 2006. ISSN 2037-416X. doi: 10.4401/ag-3141.

B. Aiazzi, L. Alparone, S. Baronti, F. Butera, E. Capobianco, L. Chiarantini, and M. Selva. Hyperspectral noise modelling and estimation. In *Proceedings of 'Hyperspectral 2010 Workshop'*, SP-683, pages 1–4, Frascati, Italy, May 2010. ESA - European Space Agency.

F. Aires, W. B. Rossow, N. A. Scott, and A. Chédin. Remote sensing from the infrared atmospheric sounding interferometer instrument 2. simultaneous retrieval of temperature, water vapor, and ozone atmospheric profiles. *Journal of Geophysical Research: Atmospheres*, 107 (D22):ACH 7–1–ACH 7–12, 2002a. ISSN 2156-2202. doi: 10.1029/2001JD001591.

F. Aires, W. B. Rossow, N. A. Scott, and A. Chédin. Remote sensing from the infrared atmospheric sounding interferometer instrument 1. compression, denoising, and first-guess retrieval algorithms. *Journal of Geophysical Research: Atmospheres*, 107(D22):ACH 6–1–ACH 6–10, 2002b. doi: 0.1029/2001JD000955.

H. Akima. A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points. *ACM Transactions on Mathematical Software (TOMS)*, 4(2):148–159, June 1978. ISSN 0098-3500. doi: 10.1145/355780.355786.

G. Allaire. *Analyse numérique et optimisation: une introduction à la modélisation mathématique et à la simulation numérique*. Éditions Ellipses, Éditions de l'École Polytechnique, first edition, 2005. ISBN 9782730212557. English translation: "Numerical Analysis and Optimization: An Introduction to Mathematical Modelling and Numerical Simulation', translated by Alan Craig.

D. R. Anderson, K. P. Burnham, and W. L. Thompson. Null hypothesis testing: problems, prevalence, and an alternative. *The journal of wildlife management*, pages 912–923, 2000.

K. Anderson, J. Dungan, and A. MacArthur. On the reproducibility of field-measured reflectance factors in the context of vegetation studies. *Remote Sensing of Environment*, 115(8):1893–1905, 2011. ISSN 0034-4257. doi: 10.1016/2011.03.012.

E. Angelopoulou, S. Lee, and R. Bajcsy. Spectral gradient: a material descriptor invariant to geometry and incident illumination. In *Proceedings of the Seventh IEEE International Conference on Computer Vision.*, volume 2, pages 861–867, September 1999. doi: 10.1109/ICCV.1999.790312.

S. Anily and A. Federgruen. Ergodicity in parametric nonstationary markov chains: An application to simulated annealing methods. *Operations Research*, 35(6):867–874, 1987. doi: 10.1287/opre.35.6.867.

A. Apan, A. Held, S. Phinn, and J. Markley. Detecting sugarcane 'orange rust' disease using EO-1 hyperion hyperspectral imagery. *International Journal of Remote Sensing*, 25(2):489–498, 2004. doi: 10.1080/01431160310001618031.

H. Arefi, P. d'Angelo, H. Mayer, and P. Reinartz. Iterative approach for efficient digital terrain model production from cartosat-1 stereo images. *Journal of Applied Remote Sensing*, 5(1): 053527–053527–19, 2011. doi: 10.1117/1.3595265.

M. A. Arostegui Jr., S. N. Kadipasaoglu, and M. K. Basheer. An empirical comparison of tabu search, simulated annealing, and genetic algorithms for facilities location problems. *International Journal of Production Economics*, 103(2):742–754, 2006. ISSN 0925-5273. doi: 10.1016/j.ijpe.2005.08.010.

R. Artstein and M. Poesio. Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34(4):555–596, 2008.

G. P. Asner, B. Braswell, D. S. Schimel, and C. A. Wessman. Ecological research needs from multiangle remote sensing data. *Remote Sensing of Environment*, 63(2):155–165, 1998. ISSN 0034-4257. doi: 10.1016/S0034-4257(97)00139-9.

I. Atkinson, F. Kamalabadi, and D. Jones. Wavelet-based hyperspectral image estimation. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS).*, volume 2, pages 743–745 vol.2, July 2003. doi: 10.1109/IGARSS.2003.1293903.

M. Bachmann, S. Adar, E. Ben-Dor, J. Biesemans, M. Grant, J. Hanus, G. Heuvelink, S. Holzwarth, A. Hueni, E. Knaeps, M. Kneubuehler, K. Meuleman, E. de Miguel, A. Pimstein, E. Prado, I. Reusen, T. Ruhtz, and M. Schaale. DJ2.2.1 - Generic Quality Indicators and Quality Layers. Technical report, EUFAR FP7, JRA2 - HYQUAPRO, March 2010a. URL www.eufar.net.

M. Bachmann, M. Habermeyer, A. Müller, R. Müller, M. Schneider, and T. Storch. Operational quality control for hyperspectral data. In *Proceedings of Hyperspectral Workshop 2010: ESA Hyperspectral Workshop, 17.*, March 2010b. ISBN 978-92-9221-247-6.

M. Bachmann, S. Adar, E. Ben-Dor, J. Biesemans, X. Briottet, M. Grant, J. Hanus, G. Heuvelink, S. Holzwarth, A. Hueni, M. Kneubuehler, K. Meuleman, E. de Miguel, I. Pérez González, I. Reusen, R. Richter, T. Ruhtz, M. Schaale, and S. Weide. DJ2.2.2 - Quality Layers for VITO, DLR, INTA, and PML. Technical report, EUFAR FP7, JRA2 - HYQUAPRO, Mai 2011a. URL www.eufar.net.

M. Bachmann, S. Adar, E. Ben-Dor, J. Biesemans, X. Briottet, M. Grant, J. Hanus, G. Heuvelink, S. Holzwarth, A. Hueni, M. Kneubuehler, K. Meuleman, E. de Miguel, I. Pérez González, I. Reusen, R. Richter, T. Ruhtz, M. Schaale, and S. Weide. DJ2.2.3 - Quality Layers for USBE, TAU-BarKal and FUB. Technical report, EUFAR FP7, JRA2 - HYQUAPRO, Mai 2011b. URL www.eufar.net.

M. Bachmann, M. Habermeyer, S. Holzwarth, G. Kerr, and I. Reusen. Presentation of the DLR's processing and archiving facility in relation to eufar. Poster, QA4EO Workshop, October 2011c.

A. Baraldi, L. Bruzzone, and P. Blonda. Quality assessment of classification and cluster maps without ground truth knowledge. *IEEE Transactions on Geoscience and Remote Sensing*, 43 (4):857–873, April 2005. ISSN 0196-2892. doi: 10.1109/TGRS.2004.843074.

I. Barrodale, R. Kuwahara, R. Poeckert, and D. Skea. Side-scan sonar image processing using thin plate splines and control point matching. *Numerical Algorithms*, 5(2):85–98, 1993. ISSN 1017-1398. doi: 10.1007/BF02212041.

S. Bates and S. McLaughlin. The estimation of stable distribution parameters. In *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics*, pages 390–394, Jul 1997. doi: 10.1109/HOST.1997.613553.

S. Bates and S. McLaughlin. The estimation of stable distribution parameters from teletraffic data. *IEEE Transactions on Signal Processing*, 48(3):865–870, Mar 2000. ISSN 1053-587X. doi: 10.1109/78.824681.

A. Baumgartner, P. Gege, C. Köhler, K. Lenhard, and T. Schwarzmaier. Characterisation methods for the hyperspectral sensor HySpex at DLR's calibration home base. In *SPIE Proceedings: Sensors, Systems, and Next-Generation Satellites XVI*, volume 8533, pages 1–8, November 2012. doi: 10.1117/12.974664.

H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. SURF: Speeded-up robust features. *Computer Vision and Image Understanding (CVIU)*, 110(3):346–359, 2008. ISSN 1077-3142. doi: 10.1016/j.cviu.2007.09.014.

A. Bayer. *Methodological Developments for Mapping Soil Constituents using Imaging Spectroscopy*. PhD thesis, Mathematisch-Naturwissenschaftlichen Fakultät der Universität Potsdam, February 2013. Available online: http://opus.kobv.de/ubp/volltexte/2013/6439/.

A. Bayer, M. Bachmann, A. Müller, and H. Kaufmann. A comparison of feature-based mlr and pls regression techniques for the prediction of three soil constituents in a degraded south african ecosystem. *Applied and Environmental Soil Science*, 2012:1–20, 2012. doi: 10.1155/2012/971252.

R. J. Beckman and G. L. Tiet jen. Maximum likelihood estimation for the beta distribution. *Journal of Statistical Computation and Simulation*, 7(3-4):253–258, 1978. doi: 10.1080/00949657808810232.

J. Beekhuizen, M. Bachmann, E. Ben-Dor, J. Biesemans, M. Grant, G. Heuvelink, A. Hueni, M. Kneubuehler, E. de Miguel, A. Pimstein, E. Prado, I. Reusen, T. Ruhtz, and M. Schaale. DJ2.1.2 - Report on Full error propagation concept. Technical report, EUFAR FP7, JRA2 - HYQUAPRO, August 2009a. URL www.eufar.net.

J. Beekhuizen, M. Bachmann, E. Ben-Dor, J. Biesemans, M. Grant, G. Heuvelink, A. Hueni, M. Kneubuehler, T. Malthus, E. de Miguel, A. Pimstein, E. Prado, I. Reusen, T. Ruhtz, and M. Schaale. DJ2.1.1 - Report on Calibration, geometric and atmospheric error model/concept. Technical report, EUFAR FP7, JRA2 - HYQUAPRO, June 2009b. URL www.eufar.net.

E. Ben-Dor, S. Adar, I. Livne, C. Fischer, G. Kerr, C. Ehrler, V. Kopačková, K. Zelenková, O. R., and S. Chevrel. Earth-observation mission planning for environmental monitoring of mining impact. Technical Report EO-Miners Deliverable 2.1, EO-Miners FP7, July 2011.

E. Ben-Dor, I. Livne, G. Notesco, S. Adar, A. Brook, G. Kopačková, Veronika Zibret, H. Coetzee, R. Oglivy, C. Fischer, C. Ehrler, and G. Kerr. Thematic accuracy and validation. Technical Report EO-Miners Deliverable 2.5, EO-Miners FP7, January 2014a.

E. Ben-Dor, I. Livne, G. Notesco, S. Adar, V. Kopačková, D. Rogge, G. Zibret, C. Ehrler, C. Fischer, and G. Kerr. Thematic accuracy and validation. Technical Report EO-Miners Deliverable 2.4, EO-Miners FP7, January 2014b.

A. Ben-Menahem and S. J. Singh. *Seismic waves and sources*. Courier Dover Publications, 2000.

E. M. Bennett, R. Alpert, and A. C. Goldstein. Communications through limited-response questioning. *Public Opinion quaterly, fall*, 1954.

A. Berk, L. Bernstein, G. Anderson, P. Acharya, D. Robertson, J. Chetwynd, and S. Adler-Golden. MODTRAN cloud and multiple scattering upgrades with application to AVIRIS. *Remote Sensing of Environment*, 65(3):367–375, 1998. ISSN 0034-4257. doi: 10.1016/S0034-4257(98)00045-5.

A. Berk, P. Acharya, G. Anderson, and B. Gossage. Recent developments in the MODTRAN®atmospheric model and implications for hyperspectral compensation. In *IEEE International conference on Geoscience and Remote Sensing Symposium, IGARSS*, volume 2, pages II–262–II–265, July 2009. doi: 10.1109/IGARSS.2009.5418058.

C. Bernard. *Wavelets and ill-posed problems: optic flow and scattered data interpolation*. PhD thesis, École Polytechnique, November 1999.

C. P. Bernard, S. G. Mallat, and J.-J. Slotine. Wavelet interpolation networks. In *European Symposium on Artificial Neural Networks (ESANN'98)*, april 1998.

J. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(2):354–379, April 2012. ISSN 1939-1404. doi: 10.1109/JSTARS.2012.2194696.

P. Bofill and M. Zibulevsky. Underdetermined blind source separation using sparse representations. *Signal Processing*, 81(11):2353–2362, 2001. ISSN 0165-1684. doi: 10.1016/S0165-1684(01)00120-7.

F. L. Bookstein. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989. ISSN 0162-8828. doi: 10.1109/34.24792.

L. Bousquet, S. Lachérade, S. Jacquemoud, and I. Moya. Leaf BRDF measurements and model for specular and diffuse components differentiation. *Remote Sensing of Environment*, 98(2–3): 201–211, 2005. ISSN 0034-4257. doi: 10.1016/j.rse.2005.07.005.

C. Brannstrom and A. M. Filippi. Remote classification of cerrado (savanna) and agricultural land covers in northeastern brazil. *Geocarto International*, 23(2):109–134, 2008. doi: 10.1080/10106040701596767.

S. Brogaard and R. Ólafsdóttir. Ground-truths or Ground-lies? Environmental sampling for remote sensing application exemplified by vegetation cover data. Technical report, Lund Electronic Reports in Physical Geography, October 1997.

A. Brook and E. Ben Dor. Spectral quality indicators for hyperspectral data. In *Proceedings of the 3^{rd} IEEE workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–5, 2011. doi: 10.1109/WHISPERS.2011.6080934.

S. P. Brooks, N. Friel, and R. King. Classical model selection via simulated annealing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):503–520, 2003. ISSN 1467-9868. doi: 10.1111/1467-9868.00399.

L. G. Brown. A survey of image registration techniques. *ACM Computing Survey*, 24(4):325–376, Dec. 1992. ISSN 0360-0300. doi: 10.1145/146370.146374.

L. Bruzzone and M. Marconcini. Toward the automatic updating of land-cover maps by a domain-adaptation svm classifier and a circular validation strategy. *IEEE Transactions on Geoscience and Remote Sensing*, 47(4):1108–1122, April 2009. ISSN 0196-2892. doi: 10.1109/TGRS.2008.2007741.

A. Buades, B. Coll, and J. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling and Simulation*, 4(2):490–530, 2005. doi: 10.1137/040616024.

M. D. Buhmann. Radial basis functions. *Acta Numerica*, 9:1–38, 1 2000. ISSN 1474-0508.

C. Cao, X. Xiong, A. Wu, and X. Wu. Assessing the consistency of AVHRR and MODIS L1B reflectance for generating fundamental climate data records. *Journal of Geophysical Research*, 113(9):D09114–1–D09114–10, may 2008. doi: 10.1029/2007JD009363.

J.-F. Cardoso. High-order contrasts for independent component analysis. *Neural Comput.*, 11 (1):157–192, Jan. 1999. ISSN 0899-7667. doi: 10.1162/089976699300016863.

M. J. Carlotto. Effect of errors in ground truth on classification accuracy. *International Journal of Remote Sensing*, 30(18):4831–4849, 2009. doi: 10.1080/01431160802672864.

J. Carpenter and J. Bithell. Bootstrap confidence intervals: when, which, what? a practical guide for medical statisticians. *Statistics in Medicine*, 19:1141–1164, 2000.

J. C. Carr, R. K. Beatson, J. B. Cherrie, T. J. Mitchell, W. R. Fright, B. C. McCallum, and T. R. Evans. Reconstruction and representation of 3d objects with radial basis functions. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, pages 67–76, New York, NY, USA, 2001. ACM. ISBN 1-58113-374-X. doi: 10.1145/383259.383266.

G. A. Carter and R. L. Miller. Early detection of plant stress by digital imaging within narrow stress-sensitive wavebands. *Remote Sensing of Environment*, 50(3):295–302, 1994. ISSN 0034-4257. doi: 10.1016/0034-4257(94)90079-5.

K. Cawse-Nicholson, M. Sears, and A. Robin. Evaluation of bands containing spectrally correlated noise in hyperspectral imagery. In *Proceedings of the* 5^{th} *IEEE workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–4, Gainesville, FL, July 2013. University of Gainesville, IEEE.

D. Cerra, R. Müller, J. Bieniarz, and P. Reinartz. On the application of spectral unmixing for noise reduction. In *Proceedings of the* 5^{th} *IEEE workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–4, Gainesville, FL, July 2013. University of Gainesville, IEEE.

T. Chan and C.-K. Wong. Total variation blind deconvolution. *IEEE Transactions on Image Processing*, 7(3):370–375, Mar 1998. ISSN 1057-7149. doi: 10.1109/83.661187.

T. F. Chan and C. Wong. Convergence of the alternating minimization algorithm for blind deconvolution. *Linear Algebra and its Applications*, 316(1-3):259–285, 2000. ISSN 0024-3795. doi: 10.1016/S0024-3795(00)00141-5. Special Issue: Conference celebrating the 60th birthday of Robert J. Plemmons.

G. Chander, J. B. Christopherson, G. L. Stensaas, and P. M. Teillet. Online catalog of world-wide test sites for the post-launch characterization and calibration of optical sensors. In *Proceedings of the* 58^{th} *International Astronomical Congress*, pages 2043–2051, Hyderabad, India, 2007. Online Catalogue available at: http://calval.cr.usgs.gov/rst-resources/sites_catalog/, last visited: Aug. 2014.

G. Chander, D. Helder, D. Aaron, N. Mishra, and A. Shrestha. Assessment of spectral, misregistration, and spatial uncertainties inherent in the cross-calibration study. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3):1282–1296, March 2013a. ISSN 0196-2892. doi: 10.1109/TGRS.2012.2228008.

G. Chander, T. Hewison, N. Fox, X. Wu, X. Xiong, and W. Blackwell. Overview of intercalibration of satellite instruments. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3): 1056–1080, March 2013b. ISSN 0196-2892. doi: 10.1109/TGRS.2012.2228654.

C.-I. Chang. Spectral information divergence for hyperspectral image analysis. In *Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, volume 1, pages 509–511 vol.1, 1999. doi: 10.1109/IGARSS.1999.773549.

C.-I. Chang. An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis. *IEEE Transactions on Information Theory*, 46(5): 1927–1932, Aug 2000. ISSN 0018-9448. doi: 10.1109/18.857802.

C.-I. Chang and Q. Du. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 42(3):608–619, March 2004. ISSN 0196-2892. doi: 10.1109/TGRS.2003.819189.

Q.-S. Chen, M. Defrise, and F. Deconinck. Symmetric phase-only matched filtering of fourier-mellin transforms for image registration and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(12):1156–1168, Dec 1994. ISSN 0162-8828. doi: 10.1109/34.387491.

Y. Chen, Z. Xue, and R. S. Blum. Theoretical analysis of an information-based quality measure for image fusion. *Information Fusion*, 9(2):161–175, 2008. ISSN 1566-2535. doi: 10.1016/j.inffus.2007.03.002.

S. Chevrel, H. Coetzee, E. Ben-Dor, C. Ehrler, E. W. Falk, C. Fischer, C. Jordan, G. Kerr, V. Kopackova, E. Kylychbaev, P. Schepelmann, S. Solar, S. Adar, and K. Zelenková. Towards sustained and accepted geo-spatial information products for mining and resources management. Presentation, 7th European Congress on Regional Geoscientific Cartography and Information Systems (EUREGEO), Bologna, Italy, June 2012.

S. C. Choi and R. Wette. Maximum likelihood estimation of the parameters of the gamma distribution and their bias. *Technometrics*, 11(4):683–690, 1969. doi: 10.1080/00401706.1969.10490731.

N. R. Chrisman. Beyond stevens: A revised approach to measurement for geographic information. In *Autocarto Confference*, pages 271–280, 1995.

N. R. Chrisman. Rethinking levels of measurement for cartography. *Cartography and Geographic Information Systems*, 25(4):231–242, 1998. doi: 10.1559/152304098782383043.

CISQ Technical Work Groups for: Reliability, Performance Efficiency, Security and Maintainability. CISQ specifications for automated quality characteristic measures. Technical Report CISQ–TR–2012–01 ©OMG 2012, CISQ - Consortium for IT Software Quality, Januray 2012.

R. Clark, G. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, and S. Sutley. Usgs digital spectral library splib06a: U.s. geological survey, digital data series 231. Technical report, U.S. Geological Survey, 2007. Available online: http://speclab.cr.usgs.gov/spectral.lib06.

R. N. Clark. *Remote Sensing for the Earth Science*, volume 3, chapter 1: Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy, pages 3–58. John Wiley and Sons, A.N. Rencz, ed., New York, 1999.

R. N. Clark and T. L. Roush. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research: Solid Earth*, 89(B7):6329–6340, 1984. ISSN 2156-2202. doi: 10.1029/JB089iB07p06329.

T. Cocks, R. Jenssen, A. Stewart, I. Wilson, and T. Shields. The HyMap^{TM} airborne hyperspectral sensor: the system, calibration and performance. In *First EARSeL workshop on imaging spectrometry*, pages 37–42, 1998.

J. Cohen. A coefficient of agreement for nominal scales. *Educational And Psychological Measurement*, 20(1):37–46, 1960. doi: 10.1177/001316446002000104.

J. Cohen. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, 70(4):213, 1968.

A. Comar, F. Baret, F. Viénot, L. Yan, and B. de Solan. Wheat leaf bidirectional reflectance measurements: Description and quantification of the volume, specular and hot-spot scattering features. *Remote Sensing of Environment*, 121(0):26–35, 2012. ISSN 0034-4257. doi: 10.1016/j.rse.2011.01.028.

R. G. Congalton. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. *Photogrammetric Engineering & Remote Sensing*, 54(5):593–600, 1988.

R. G. Congalton. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1):35–46, 1991. ISSN 0034-4257. doi: 10.1016/0034-4257(91)90048-B.

R. G. Congalton. Accuracy assessment and validation of remotely sensed and other spatial information. *International Journal of Wildland Fire*, 10(4):321–328, 2001. doi: 10.1071/WF01031.

B. R. Corner, R. M. Narayanan, and S. E. Reichenbach. Noise estimation in remote sensing imagery using data masking. *International Journal of Remote Sensing*, 24(4):689–702, 2003. doi: 10.1080/01431160210164271.

N. Cressie. The origins of kriging. *Mathematical Geology*, 22(3):239–252, 1990. ISSN 0882-8121. doi: 10.1007/BF00889887.

B. Datt. A new reflectance index for remote sensing of chlorophyll content in higher plants: Tests using eucalyptus leaves. *Journal of Plant Physiology*, 154(1):30–36, 1999. ISSN 0176-1617. doi: 10.1016/S0176-1617(99)80314-9.

I. Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41(7):909–996, 1988. ISSN 1097-0312. doi: 10.1002/cpa.3160410705.

I. Daubechies, I. Guskov, P. Schröder, and W. Sweldens. Wavelets on irregular point sets. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357(1760):2397–2413, 1999. doi: 10.1098/rsta.1999.0439.

R. Davidson and J. G. MacKinnon. Bootstrap Tests: How Many Bootstraps? Technical Report 1036, Queen's University, Department of Economics, March 2001.

O. A. De Carvalho and P. Meneses. Spectral correlation mapper (SCM): An improvement on the spectral angle mapper (SAM). In *Summaries of the nineth JPL Airborne Earth Science Workshop*, volume 18, pages 1–9. Jet Propulsion Laboratories, JPL Publication, 2000.

E. De Castro and C. Morandi. Registration of translated and rotated images using finite fourier transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(5):700–703, Sept 1987. ISSN 0162-8828. doi: 10.1109/TPAMI.1987.4767966.

C. Dean, I.-L. Chang, Z. Li, M. Weinreb, X. Wu, and F. Yu. Recent advances in calibration of the GOES imager visible channel at NOAA. In X. G. e. James J. Butler, Xiaoxiong Xiong, editor, *Proceedings of SPIE, Earth Observing Systems*, volume 8510-XVII, pages 85100M–85100M–11, October 2012. doi: 10.1117/12.929036.

E. Dearing, P. Staes, and T. Prorok, editors. *CAF works - better service for the citizens by using CAF*. Publisher, 2006. Available online: http://www.eipa.eu/en/topic/show/&tid=191#.

J.-M. Delvit, D. Leger, S. Roques, C. Valorge, and F. Viallefont-Robinet. Signal to noise ratio assessment from nonspecific views. In *Proceedings of SPIE*, volume 4541, pages 370–381, 2002. doi: 10.1117/12.454171.

A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society, Series B*, 1977.

P. E. Dennison, K. Q. Halligan, and D. A. Roberts. A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper. *Remote Sensing of Environment*, 93(3):359 – 367, 2004. ISSN 0034-4257. doi: 10.1016/j.rse.2004.07.013.

S. Descamps, X. Descombes, A. Bechet, and J. Zerubia. Automatic flamingo detection using a multiple birth and death process. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1113–1116, March 2008. doi: 10.1109/ICASSP.2008.4517809.

G. Deslauriers and S. Dubuc. Symmetric iterative interpolation processes. *Constructive Approximation*, 5:49–68, 1989. ISSN 0176-4276. 10.1007/BF01889598.

T. J. DiCiccio and B. Efron. More accurate confidence intervals in exponential families. Technical Report 368, Department of Statistics, Stanford University, February 1991.

T. J. DiCiccio and B. Efron. Bootstrap confidence intervals. *Statistical Science*, 11(3):189–228, 1996. doi: 10.1214/ss/1032280214.

M. Do and M. Vetterli. Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance. *IEEE Transactions on Image Processing*, 11(2):146–158, 2002. ISSN 1057-7149. doi: 10.1109/83.982822.

D. Doelling, D. Morstad, B. Scarino, R. Bhatt, and A. Gopalan. The characterization of deep convective clouds as an invariant calibration target and as a visible calibration technique. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3):1147–1159, March 2013. ISSN 0196-2892. doi: 10.1109/TGRS.2012.2225066.

W. Dorigo. *Retrieving canopy variables by radiative transfer model inversion - a regional approach for imaging spectrometer data.* PhD thesis, Faculty Wissenschaftszentrum Weihenstephan,Technical University of Munich, 2008.

J. Duchon. Splines minimizing rotation-invariant semi-norms in sobolev spaces. In W. Schempp and K. Zeller, editors, *Constructive Theory of Functions of Several Variables*, volume 571 of *Lecture Notes in Mathematics*, pages 85–100. Springer Berlin Heidelberg, 1977. ISBN 978-3-540-08069-5. doi: 10.1007/BFb0086566.

E. Dumic, S. Grgic, and M. Grgic. New image-quality measure based on wavelets. *Journal of Electronic Imaging*, 19(1):011–018, July 2010.

ECSS-E-HB-40A Working Group. *Space engineering, Software engineering handbook.* ESA Requirements and Standards Division, 2013.

B. Efron. Better bootstrap confidence intervals. Technical Report 14, Laboratory for computational statistics - Department of Statistics - Stanford University, November 1984.

C. Ehrler. *Scale-Wavelength Decomposition of Hyperspectral Signals: Use for Mineral Classification & Quantification.* PhD thesis, Karlsruher Instituts für Technologie (KIT), 2014.

C. Ehrler, C. Fischer, G. Kerr, and D. Rogge. Raw and radiometric data quality indices pre and post correction. Technical Report EO-Miners Deliverable 2.2, EO-Miners FP7, October 2013.

EO-Miners. Eo-miners website, 2014. http://www.eo-miners.eu/.

EUFAR - EUropean Facility for Airborne Research. Eufar website, 2014. http://www.eufar.net/.

European Parliament, Council. INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO-19115 and EN ISO-19119. Technical report, European Commission Joint Research Centre, 2013.

EXELIS. Envi-tutorials: Open and browse data. Technical report, Exelis Visual Information Solution, 2014. Available online: http://www.exelisvis.com/docs/OpenBrowseData.html.

U. T. Eysel. Illusions and perceived images in the primate brain. *Science*, 302(5646):789–791, 2003. doi: 10.1126/science.1091065.

FAA. Risk analysis (RASEM) correspondence course. Technical report, Federal Aviation Agency - FAA, 2014. Available online: http://www.faa.gov/aircraft/air_cert/design_approvals/csta/publications/.

E. W. Falck, I. Livne, S. Adar, E. Ben-Dor, H. Coetzee, C. Ehrler, C. Fischer, G. Kerr, F. C., C. J. Jordan, V. Kopačková, K. Zelenková, O. R., and S. Chevrel. Status quo of monitoring environmental candidate indicators by earth-observation services. Technical Report EO-Miners Deliverable 1.8, EO-Miners FP7, October 2013.

R. Falk and C. W. Greenbaum. Significance tests die hard: The amazing persistence of a probabilistic misconception. *Theory & Psychology*, 5(1):75–98, 1995. doi: 10.1177/0959354395051004.

E. F. Fama and R. Roll. Some properties of symmetric stable distributions. *Journal of the American Statistical Association*, 63(323):817–836, September 1968.

E. F. Fama and R. Roll. Parameter estimates for symmetric stable distributions. *Journal of the American Statistical Association*, 66(334):331–338, 1971. doi: 10.1080/01621459.1971.10482264.

P. Fan, A. Men, M. Chen, and B. Yang. Color-surf: A surf descriptor with local kernel color histograms. In *Proceedings of the IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC)*., pages 726–730, Nov 2009. doi: 10.1109/ICNIDC.2009.5360809.

T. G. Farr, P. A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, D. Seal, S. Shaffer, J. Shimada, J. Umland, M. Werner, M. Oskin, D. Burbank, and D. Alsdorf. The shuttle radar topography mission. *Reviews of Geophysics*, 45, 2007. doi: 10.1029/2005RG000183.

M. Farzam and S. Beheshti. Robust hyperspectral signal unmixing in the presence of correlated noise. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*., pages 1361–1364, May 2011. doi: 10.1109/ICASSP.2011.5946665.

A. Faul and M. Powell. Proof of convergence of an iterative technique for thin plate spline interpolation in two dimensions. *Advances in Computational Mathematics*, 11(2-3):183–192, 1999. ISSN 1019-7168. doi: 10.1023/A:1018923925800.

H. G. Feichtinger and K. Gröchenig. Theory and practice of irregular sampling. In *Wavelets: mathematics and applications*, volume 1994, pages 305–363. CRC Press, Boca Raton, FL, 1994. ISBN ISBN=0-8493-8271-8/hbk.

G. C. Feng. Factors affecting intercoder reliability: a monte carlo experiment. *Quality & Quantity*, 47(5):2959–2982, 2012. ISSN 0033-5177. doi: 10.1007/s11135-012-9745-9.

T. S. Ferguson. Maximum likelihood estimates of the parameters of the cauchy distribution for samples of size 3 and 4. *Journal of the American Statistical Association*, 73(361):211–213, March 1978. doi: 10.1080/01621459.1978.10480031.

A. Fernández-Renau, J. A. Gómez, and E. de Miguel. The INTA AHS system. In *Sensors, Systems, and Next-Generation Satellites*, volume 9, pages 59781–59788, October 2005. doi: 10.1117/12.629440.

A. Filippi and J. Jensen. Effect of continuum removal on hyperspectral coastal vegetation classification using a fuzzy learning vector quantizer. *IEEE Transactions on Geoscience and Remote Sensing*, 45(6):1857–1869, June 2007. ISSN 0196-2892. doi: 10.1109/TGRS.2007.894929.

C. Fischer and C. Ehrler. Report on In-flight Calibration and Atmospheric Correction of DeMinTIR Imagery. Technical report, German Aerospace Center - Entity airborne remote sensing, 2012.

C. Fischer, C. Ehrler, G. Kerr, and D. Rogge. Presentation of algorithms. Technical Report EO-Miners Deliverable 3.3, EO-Miners FP7, January 2014a.

C. Fischer, C. Ehrler, G. Kerr, D. Rogge, C. Jordan, V. Kopačková, K. Zelenková, J. Misurec, I. Livne, S. Adar, E. Ben-Dor, H. Coetzee, D. Van Tonder, G. Zibret, R. Sajn, S. Chevrel, A. Bourguignon, E. Kylychbaerv, U. Abdybachaev, E. Mambetakiyev, and E. Dudaschvili. Data acquisition and validation. Technical Report EO-Miners Deliverable 3.2, EO-Miners FP7, January 2014b.

C. Fischer, G. Kerr, C. Ehrler, D. Rogge, E. Ben-Dor, I. Livne, G. Notesco, and A. Brook. Atmospheric correction protocol and validation results. Technical Report EO-Miners Deliverable 2.3, EO-Miners FP7, January 2014c.

J. L. Fleiss. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76 (5):378–382, 1971.

J. L. Fleiss and J. Cohen. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, 33:613–619, 1973.

J. L. Fleiss, B. Levin, and M. Cho Paik. *Statistical Methods for Rates and Proportions*, chapter 18, The Measurement of Interrater Agreement, pages 598–626. John Wiley & Sons, Inc., third edition, 2003. ISBN 0-471-52629-0.

G. M. Foody. Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data. *International Journal of Remote Sensing*, 17(7):1317–1340, 1996. doi: 10.1080/01431169608948706.

G. M. Foody. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1):185–201, 2002. ISSN 0034-4257. doi: 10.1016/S0034-4257(01)00295-4.

G. M. Foody. Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sensing of Environment*, 114(10):2271–2285, 2010. ISSN 0034-4257. doi: 10.1016/j.rse.2010.05.003.

G. M. Foody, L. See, S. Fritz, M. Van der Velde, C. Perger, C. Schill, and D. S. Boyd. Assessing the accuracy of volunteered geographic information arising from multiple contributors to an internet based collaborative project. *Transactions in GIS*, 17(6):847–860, 2013. doi: 10.1111/tgis.12033.

B. Fougnie, G. Bracco, B. Lafrance, C. Ruffel, O. Hagolle, and C. Tinel. PARASOL in-flight calibration and performance. *Appl. Opt.*, 46(22):5435–5451, Aug 2007. doi: 10.1364/AO.46.005435.

R. Franke. Smooth interpolation of scattered data by local thin plate splines. *Computers & Mathematics with Applications*, 8(4):273–281, 1982. ISSN 0898-1221. doi: 10.1016/0898-1221(82)90009-8.

M. Frigo and S. G. Johnson. The design and implementation of FFTW3. *Proceedings of the IEEE*, 93(2):216–231, Feb 2005. ISSN 0018-9219. doi: 10.1109/JPROC.2004.840301.

M. Galassi and et al. *GNU Scientific Library Reference Manual (3rd Ed.)*. Network Theory Ltd, January 2012. ISBN 0954612078. Available online: http://www.gnu.org/software/gsl/.

J. Gamon, J. Peñuelas, and C. Field. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment*, 41(1):35–44, 1992. ISSN 0034-4257. doi: 10.1016/0034-4257(92)90059-S.

B.-C. Gao. An operational method for estimating signal to noise ratios from data acquired with imaging spectrometers. *Remote Sensing of Environment*, 43(1):23–33, 1993. ISSN 0034-4257. doi: 10.1016/0034-4257(93)90061-2.

B.-c. Gao. NDWI: A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3):257–266, 1996. ISSN 0034-4257. doi: 10.1016/S0034-4257(96)00067-3.

P. Garcia-Allende, O. Conde, J. Mirapeix, A. Cubillas, and J. Lopez-Higuera. Data processing method applying principal component analysis and spectral angle mapper for imaging spectroscopic sensors. *IEEE Sensors Journal*, 8(7):1310–1316, July 2008. ISSN 1530-437X. doi: 10.1109/JSEN.2008.926923.

P. Garg. Genetic algorithms, tabu search and simulated annealing: a comparison between three approaches for the cryptanalisys of transposition cipher. *Journal of Theoretical and Applied Information Technology*, 2009.

P. Gege, J. Fries, P. Haschberger, P. Schötz, H. Schwarzer, P. Strobl, B. Suhr, and V. W. Chb handbook. online, September 2008.

P. Gege, J. Fries, P. Haschberger, P. Schötz, H. Schwarzer, P. Strobl, B. Suhr, G. Ulbrich, and W. J. Vreeling. Calibration facility for airborne imaging spectrometers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(4):387–397, 2009. ISSN 0924-2716. doi: 10.1016/j.isprsjprs.2009.01.006.

J. Geronimo, D. Hardin, and P. Massopust. Fractal functions and wavelet expansions based on several scaling functions. *Journal of Approximation Theory*, 78(3):373–401, 1994. ISSN 0021-9045. doi: 10.1006/jath.1994.1085.

D. R. Gerwe, C. E. Luna, and B. Calef. Information theoretic based image quality evaluation. In *Frontiers in Optics 2009/Laser Science XXV/Fall 2009 OSA Optics & Photonics Technical Digest*, pages 0–0. Optical Society of America, 2009. doi: 10.1364/SRS.2009.STuC1.

J.-M. Geusebroek, R. Van den Boomgaard, A. W. M. Smeulders, and H. Geerts. Color invariance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(12):1338–1350, Dec 2001. ISSN 0162-8828. doi: 10.1109/34.977559.

B. Gidas. Nonstationary markov chains and convergence of the annealing algorithm. *Journal of Statistical Physics*, 39(1-2):73–131, 1985. ISSN 0022-4715. doi: 10.1007/BF01007975.

G. Gigerenzer. Mindless statistics. *Journal of Behavioral and Experimental Economics (formerly The Journal of Socio-Economics)*, 33(5):587–606, 2004.

A. F. Goetz. Three decades of hyperspectral remote sensing of the earth: A personal view. *Remote Sensing of Environment*, 113, Supplement 1(0):5–16, 2009. ISSN 0034-4257. doi: 10.1016/j.rse.2007.12.014. Imaging Spectroscopy Special Issue.

A. Goodman, Leo and W. H. Kruskal. Measure of association for cross classifications. *Journal of the American Statistical Association*, 49(268):732–764, December 1954.

A. Goodman, Leo and W. H. Kruskal. Measure of association for cross classifications ii: Further discussion and references. *Journal of the American Statistical Association*, 54(285):123–163, March 1959.

A. Goshtasby. Piecewise cubic mapping functions for image registration. *Pattern Recogn.*, 20(5): 525–533, Oct. 1987. ISSN 0031-3203. doi: 10.1016/0031-3203(87)90079-3.

A. Goshtasby. Image registration by local approximation methods. *Image and Vision Computing*, 6(4):255–261, 1988. ISSN 0262-8856. doi: 10.1016/0262-8856(88)90016-9.

A. Goshtasby and G. Stockman. Point pattern matching using convex hull edges. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-15(5):631–637, Sept 1985. ISSN 0018-9472. doi: 10.1109/TSMC.1985.6313439.

M. Govender, K. Chetty, and H. Bulcock. A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water SA*, 33(2):145–152, 2007. ISSN 0378-4738. doi: 10.4314/wsa.v33i2.49049.

M. Grabowski and K. Roberts. Human and organizational error in large scale systems. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 26(1):2–16, Jan 1996. ISSN 1083-4427. doi: 10.1109/3468.477856.

J. Grace, C. Nichol, M. Disney, P. Lewis, T. Quaife, and P. Bowyer. Can we measure terrestrial photosynthesis from space directly, using spectral reflectance and fluorescence? *Global Change Biology*, 13(7):1484–1497, 2007. ISSN 1365-2486. doi: 10.1111/j.1365-2486.2007.01352.x.

S. Graham. Earth observing-1. Technical report, NASA - Earth Observatory, November 2000. URL http://earthobservatory.nasa.gov/Features/EO1/eo1_2.php. Last visited: July 2014.

V. Granville, M. Krivánek, and J. P. Rasson. Simulated annealing: A proof of convergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6):652–656, June 1994. ISSN 0162-8828. doi: 10.1109/34.295910.

A. Green, M. Berman, P. Switzer, and M. Craig. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26(1):65–74, Jan 1988. ISSN 0196-2892. doi: 10.1109/36.3001.

R. Green, B. Pavri, and T. Chrien. On-orbit radiometric and spectral calibration characteristics of eo-1 hyperion derived with an underflight of AVIRIS and in situ measurements at salar de arizaro, argentina. *IEEE Transactions on Geoscience and Remote Sensing*, 41(6):1194–1203, 2003. ISSN 0196-2892. doi: 10.1109/TGRS.2003.813204.

R. Gupta. *Remote Sensing Geology*. Springer, second edition, 2003. ISBN 9783540431855.

G. Guyot, F. Baret, and D. J. Major. High spectral resolution: Determination of specral shifts between the red and the near infrared. *International Archives of Photogrammetry and Remote Sensing*, 11:750–760, July 1988.

K. L. Gwet. Kappa statistic is not satisfactory for assessing the extent of agreement between raters. *Statistical Methods For Inter-Rater Reliability Assessment*, 1(1):1–5, 2002.

M. Habermeyer, M. Bachmann, S. Holzwarth, R. Müller, and R. Richter. Incorporating a push-broom scanner into a generic hyperspectral processing chain. In *Proceedings of the IEEE InternationalGeoscience and Remote Sensing Symposium (IGARSS)*, pages 5293–5296, July 2012. doi: 10.1109/IGARSS.2012.6352414.

O. Hagolle, J.-M. Nicolas, B. Fougnie, F. Cabot, and P. Henry. Absolute calibration of VEGETATION derived from an interband method based on the sun glint over ocean. *IEEE Transactions on Geoscience and Remote Sensing*, 42(7):1472–1481, July 2004. ISSN 0196-2892. doi: 10.1109/TGRS.2004.826805.

B. Hajek. Cooling schedules for optimal annealing. *Mathematics of Operations Research*, 1988.

S.-H. Ham and B. J. Sohn. Assessment of the calibration performance of satellite visible channels using cloud targets: application to meteosat-8/9 and mtsat-1r. *Atmospheric Chemistry and Physics*, 10(22):11131–11149, 2010. doi: 10.5194/acp-10-11131-2010.

R. L. Hardy. Multiquadric equations of topography and other irregular surfaces. *Journal of Geophysical Research*, 76(8):1905–1915, 1971. ISSN 2156-2202. doi: 10.1029/JB076i008p01905.

H. L. Harter and A. H. Moore. Maximum-likelihood estimation of the parameters of gamma and weibull populations from complete and from censored samples. *Technometrics*, 7(4):639–643, 1965. doi: 10.1080/00401706.1965.10490304.

R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004. ISBN 0521540518.

W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. doi: 10.1093/biomet/57.1.97.

A. F. Hayes and K. Krippendorff. Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1):77–89, 2007.

A. K. Heidinger, C. Cao, and J. T. Sullivan. Using moderate resolution imaging spectrometer (modis) to calibrate advanced very high resolution radiometer reflectance channels. *Journal of Geophysical Research: Atmospheres*, 107(D23):AAC 11–1–AAC 11–10, 2002. ISSN 2156-2202. doi: 10.1029/2001JD002035.

P. Henry, G. Chander, B. Fougnie, C. Thomas, and X. Xiong. Assessment of spectral band impact on intercalibration over desert sites using simulation based on eo-1 hyperion data. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3):1297–1308, March 2013. ISSN 0196-2892. doi: 10.1109/TGRS.2012.2228210.

T. Hesterberg, D. S. Moore, S. Monaghan, A. Clipson, R. Epstein, and B. A. Craig. *Introduction to the Practice of Statistics*, chapter 16: Bootstrap Methods and Permutation Tests. Ruth Baruth, 6 edition, 2007.

T. Hilker, N. C. Coops, F. G. Hall, T. A. Black, M. A. Wulder, Z. Nesic, and P. Krishnan. Separating physiologically and directionally induced changes in PRI using BRDF models. *Remote Sensing of Environment*, 112(6):2777–2788, 2008. ISSN 0034-4257. doi: 10.1016/j.rse.2008.01.011.

S. Holzwarth and M. Freer. Common metadata for hyperspectral image data. Technical report, EUFAR, N6SP-Standards and Protocols, 2011. URL www.eufar.net.

M. Hosny. Comparing genetic algorithms and simulated annealing for solving the pickup and delivery problem with time windows. In *In: Proceedings of the 2011 International Conference on Artificial Intelligence, ICAI*, volume 11, pages 513–519, 2011.

R. Höver, M. Di Luca, G. Szekely, and M. Harders. Computationally efficient techniques for data-driven haptic rendering. In *EuroHaptics conference, 2009 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2009. Third Joint*, pages 39–44, 2009. doi: 10.1109/WHC.2009.4810814.

L. M. Hsu and R. Field. Interrater agreement measures: Comments on κ_n, Cohen's κ, Scott's π, and Aickin's α. *Understanding Statistics*, 2(3):205–219, 2003. doi: 10.1207/S15328031US0203_03.

L. Ingber. Simulated annealing: Practice versus theory. *Math. Comput. Model.*, 18(11):29–57, Dec. 1993. ISSN 0895-7177. doi: 10.1016/0895-7177(93)90204-C.

IRARS: Imagery Resolution Assessments and Reporting Standards Committee. National image interpretability rating scales (NIIRS), 2012. Available online: http://www.fas.org/irp/imint/niirs.htm.

A. Iske. Reconstruction of smooth signals from irregular samples by using radial basis function approximation. In Y. Lyubarskii, editor, *Proceedings of the 1999 International workshop on sampling theory and applications*, pages 82–87. The Norwegian University of science and Technology, 1999.

H. Jahn and R. Reulke. A sensor-based approach to image quality. *Photogrammetrie-Fernerkundung-Geoinformation*, 2012(1):19–27, 2012.

A. Jalobeanu, L. Blanc-Féraud, and J. Zerubia. Estimation of blur and noise parameters in remote sensing. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 4, pages IV–3580–IV–3583, May 2002. doi: 10.1109/ICASSP.2002.5745429.

A. Janicki and A. Weron. Can one see α-stable variables and processes? *Statistical Science*, 9 (1):109–126, 02 1994. doi: 10.1214/ss/1177010656.

A. Janssen and T. Pauls. How do boostrap and permutation tests work? *The Annals of Statistics*, 31(3):768–806, 2003.

JCGM, Joint Committee for Guides in Metrology. Evaluation of Measurement Data - Supplement 1 to the "Guide to the Expression of Uncertainty in Measurement". Technical report, BIPM - Bureau International des Poids et Mesures, 2008a.

JCGM, Joint Committee for Guides in Metrology. Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement. Technical report, BIPM - Bureau International des Poids et Mesures, September 2008b.

JCGM: Joint Committee for Guides in Metrology. Evaluation of Measurement Data - An Introduction to the "Guide to the Expression of Uncertainty in Measurement". Technical report, BIPM - Bureau International des Poids et Mesures, July 2009.

JCGM, Joint Committee for Guides in Metrology. Evaluation of Measurement Data - Supplement 2 to the "Guide to the Expression of Uncertainty in Measurement". Technical report, BIPM - Bureau International des Poids et Mesures, 2011.

JCGM, Joint Committee for Guides in Metrology. *International vocabulary of metrology – Basic and general concepts and associated terms (VIM)*. BIPM-Bureau International des Poids et Mesures, 3rd edition edition, 2012.

E. W. Johnson and J. Ross. Quantifying error in aerial survey data. *Australian Forestry*, 71(3): 216–222, 2008. doi: 10.1080/00049158.2008.10675038.

P. E. Johnson, M. O. Smith, and J. B. Adams. Simple algorithms for remote determination of mineral abundances and particle sizes from reflectance spectra. *Journal of Geophysical Research: Planets*, 97(E2):2649–2657, 1992. ISSN 2156-2202. doi: 10.1029/91JE02504.

S. G. Johnson. *The NLopt nonlinear-optimization package*. http://ab-initio.mit.edu/nlopt, January 2012.

L. N. Joppa, G. McInerny, R. Harper, L. Salido, K. Takeda, K. O'Hara, D. Gavaghan, and S. Emmott. Troubling trends in scientific software use. *Science*, 340(6134):814–815, 2013. doi: 10.1126/science.1231535.

C. Jordan. Derivation of leaf area index from quality of light on the forest floor. *Ecology*, 50: 663–666, 1969.

C. J. Jordan, S. Chevrel, H. Coetzee, E. Ben-Dor, C. Ehrler, C. Fischer, S. R. Grebby, G. Kerr, I. Livne, V. Kopacková, E. Kylychbaev, F. McEvoy, and S. Adar. EO-MINERS: Monitoring the environmental and societal impact of the extractive industry using earth observation. In IEEE, editor, *Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, pages 1741–1744, Melbourne, AU, July 2013.

S. H. Joshi, E. Klassen, A. Srivastava, and I. Jermyn. A novel representation for riemannian analysis of elastic curves in $\mathbb{R}^n$. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–7, june 2007. doi: 10.1109/CVPR.2007.383185.

A. Jourjine, S. Rickard, and O. Yilmaz. Blind separation of disjoint orthogonal signals: demixing n sources from 2 mixtures. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).*, volume 5, pages 2985–2988, 2000. doi: 10.1109/ICASSP.2000.861162.

M. G. Kendall and B. B. Smith. The problem of m rankings. *The Annals of Mathematical Statistics*, 10(3):275–287, 09 1939. doi: 10.1214/aoms/1177732186.

G. Kerr and C. Fisher. Corresponding pixels automated matcher (CPAM) and application to analysis of hyperspectral data variability. In *Proceedings of the* 5^{th} *IEEE workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–4, Gainesville FL, July 2013. University of Gainesville, IEEE.

G. Kerr, C. Fischer, M. Bachmann, and H. Stefanie. Quality assurance framework for hyperspectral data products - a case study for mining environment. Poster, QA4EO Workshop, October 2011a.

G. Kerr, S. Kurtek, and A. Srivastava. A joint model for boundaries of multiple anatomical parts. In *Proc. SPIE Medical Imaging 2011*, volume 7962, pages 796246–796246–6, March 2011b. doi: 10.1117/12.877694.

G. Kerr, M. Freer, and M. Bachmann. Testing and validation report. Technical Report N6SP-DN6.3.4, EUFAR FP7, 2012.

G. Kerr, C. Fischer, D. Klein, and E. Stein. TET-1: Report on data acquisition and validation. Technical report, German Aerospace Center, German Data Center (DLR-DFD), Oberpfaffenhofen, 82234 Wessling, Germany, May 2014.

N. Keshava and J. Mustard. Spectral unmixing. *IEEE Signal Processing Magazine*, 19(1):44–57, Jan 2002. ISSN 1053-5888. doi: 10.1109/79.974727.

P. Kessler, B. Killough, S. Gowda, B. Williams, G. Chander, and M. Qu. Ceos visualization environment (cove) tool for intercalibration of satellite instruments. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3):1081–1087, March 2013. ISSN 0196-2892. doi: 10.1109/TGRS.2012.2235841.

H. H. Kieffer. Photometric stability of the lunar surface. *Icarus*, 130(2):323–327, 1997. ISSN 0019-1035. doi: 10.1006/icar.1997.5822.

D.-O. Kim and R.-H. Park. Image quality measure using the phase quantization code. *IEEE Transactions on Consumer Electronics*, 56(2):937–945, may 2010. ISSN 0098-3063. doi: 10.1109/TCE.2010.5506023.

S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

E. Klassen, A. Srivastava, W. Mio, and S. H. Joshi. Analysis of planar shapes using geodesic paths on shape spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26 (3):372–383, 2004.

S. Kogon and D. Williams. On the characterization of impulsive noise with α-stable distributions using fourier techniques. In *Signals, Systems and Computers, 1995. 1995 Conference Record of the Twenty-Ninth Asilomar Conference on*, volume 2, pages 787–791 vol.2, Oct 1995. doi: 10.1109/ACSSC.1995.540808.

K. Kölbig and B. Schorr. A program package for the landau distribution. *Computer Physics Communications*, 31(1):97–111, 1984. ISSN 0010-4655. doi: 10.1016/0010-4655(84)90085-7.

I. A. Koutrouvelis. An iterative procedure for the estimation of the parameters of stable laws. *Communications in Statistics - Simulation and Computation*, 10(1):17–28, 1981. doi: 10.1080/03610918108812189.

H. J. Kramer, editor. *Oservation of the Earth and Its Environment: Survey of Missions and Sensors*. Springer Science & Business Media, fourth edition, 2002. ISBN 9783540423881.

H. J. Kramer. Worldview-2. Technical report, EO-Portal, 2014. Available online: https://directory.eoportal.org/web/eoportal/satellite-missions/v-w-x-y-z/worldview-2#foot24%29.

K. Krippendorff. Estimating the reliability, systematic error and random error of interval data. *Educational and Psychological Measurement*, 1970.

K. Krippendorff. Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, 30(3):411–433, 2004.

K. Krippendorff. Algorithm for bootstrapping a distribution of c_α. Technical report, University of Pennsylvania, August 2006. Updated October 2013.

K. Krippendorff. Computing Krippendorff's α-reliability. Technical report, University of Pennsylvania - Annenberg School for Communication, 2011a. Available online: http://repository.upenn.edu/asc papers/43.

K. Krippendorff. Agreement and information in the reliability of coding. *Communication Methods and Measures*, 5(2):1–20, 2011b.

D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 233–240, June 2011. doi: 10.1109/CVPR.2011.5995521.

F. Kruse, A. Lefkoff, J. Boardman, K. Heidebrecht, A. Shapiro, P. Barloon, and A. Goetz. The spectral image processing system (SIPS): interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44(2-3):145–163, 1993. ISSN 0034-4257. Airbone Imaging Spectrometry.

D. Kundur and D. Hatzinakos. Blind image deconvolution. *IEEE Signal Processing Magazine*, 13(3):43–64, May 1996. ISSN 1053-5888. doi: 10.1109/79.489268.

E. E. Kuruoğlu. Density parameter estimation of skewed α-stable distributions. *Trans. Sig. Proc.*, 49(10):2192–2201, Oct. 2001. ISSN 1053-587X. doi: 10.1109/78.950775.

S. Lacherade, B. Fougnie, P. Henry, and P. Gamet. Cross calibration over desert sites: Description, methodology, and operational implementation. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3):1098–1113, March 2013. ISSN 0196-2892. doi: 10.1109/TGRS.2012.2227061.

R. G. Laha. Review: V. m. zolotarev, one-dimensional stable distributions. *Bulletin (New Series) of the American Mathematical Society*, 20(2):270–277, 04 1989. ISSN 0273-0979. Available online: http://projecteuclid.org/euclid.bams/1183555037.

J. Lahtinen, P. Myllymäki, S. Tomi, and H. Tirri. Empirical comparison of stochastic algorithms. In J. Alander, editor, *Proceedings of the Second Nordic Workshop on Genetic Algorithms and their Applications*, pages 45–49, Vaasa, August 1996. University of Vaasa and the Finnish Artificial Intelligence Society.

J. R. Landis and G. G. Koch. The measurement of observer agreement for categorical data. *biometrics*, 33(1):159–174, 1977. doi: 10.2307/2529310.

J. C. Leachtenauer, W. Malila, J. Irvine, L. Colburn, and N. Salvaggio. General image-quality equation: GIQE. *Appl. Opt.*, 36(32):8322–8328, Nov 1997. doi: 10.1364/AO.36.008322.

C. Lenart, P. Burai, A. Smailbegovic, T. Biro, Z. Katona, and R. Andricevic. Multi-sensor integration and mapping strategies for the detection and remediation of the red mud spill in kolontar, hungary: Estimating the thickness of the spill layer using hyperspectral imaging and lidar. In *Proceedings of the* 3^{rd} *IEEE workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–4, June 2011. doi: 10.1109/WHISPERS.2011.6080917.

K. Lenhard. Determination of combined measurement uncertainty via monte carlo analysis for the imaging spectrometer ROSIS. *Applied Optics*, 51(18):4065–4072, Jun 2012. doi: 10.1364/AO.51.004065.

D. Letexier and S. Bourennane. Noise removal from hyperspectral images by multidimensional filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 46(7):2061–2069, July 2008.

A. Levin, Y. Weiss, F. Durand, and W. Freeman. Understanding and evaluating blind-deconvolution algorithms. Technical Report MIT-CSAIL-TR-2009-014, MIT: Massachusetts Institute of Technology - Computer Science and Artificial Intelligence Laboratory, Ma 02139, USA, March 2009.

A. Levin, Y. Weiss, F. Durand, and W. Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2657–2664, June 2011. doi: 10.1109/CVPR.2011.5995308.

S. Z. Li, J. Kittler, and M. Petrou. Matching and recognition of road networks from aerial images. In G. Sandini, editor, *Computer Vision — ECCV'92*, volume 588 of *Lecture Notes in Computer Science*, pages 857–861. Springer Berlin Heidelberg, 1992. ISBN 978-3-540-55426-4. doi: 10.1007/3-540-55426-2_99.

M. Ligas and M. Kulczycki. Simple spatial prediction-least squares prediction, simple kriging, and conditional expectation of normal vector. *Geodesy and Cartography*, 59(2):69–81, 2010.

K. Lins. Requirements analysis results for land cover and land use data. Technical report, US Geological Survey, March 1994.

I. Lira and W. Woeger. Comparison between the conventional and bayesian approaches to evaluate measurement data. *Metrologia*, 43(4):249–259, 2006.

H. Liu and Q. Zhou. Accuracy analysis of remote sensing change detection by rule-based rationality evaluation with post-classification comparison. *International Journal of Remote Sensing*, 25(5):1037–1050, 2004. doi: 10.1080/0143116031000150004.

N. G. Loeb. In-flight calibration of noaa avhrr visible and near-ir bands over greenland and antarctica. *International Journal of Remote Sensing*, 18(3):477–490, 1997. doi: 10.1080/014311697218908.

S. Los, P. North, W. Grey, and M. Barnsley. A method to convert AVHRR normalized difference vegetation index time series to a standard viewing and illumination geometry. *Remote Sensing of Environment*, 99(4):400–411, 2005. ISSN 0034-4257. doi: 10.1016/j.rse.2005.08.017.

D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–110, 2004. ISSN 0920-5691. 10.1023/B:VISI.0000029664.99615.94.

D. Lu, M. Batistella, E. E. de Miranda, and E. Moran. A comparative study of landsat tm and spot hrg images for vegetation classification in the brazilian amazon. *Photogramm Eng Remote Sensing*, 74(3):311–321, September 2008.

L. Lucy. An iterative technique for the rectification of observed distributions. *Astronomical Journal*, 79(6):745–754, 1974. doi: 10.1086/111605.

Z. Ma and R. L. Redmond. Tau coefficients for accuracy assessment of classification of remote sensing data. *Photogrammetric Engineering and Remote Sensing*, 61(4):435–439, 1995.

H. Maître and Y. Wu. Improving dynamic programming to solve image registration. *Pattern Recognition*, 20(4):443–462, 1987. ISSN 0031-3203. doi: 10.1016/0031-3203(87)90071-9.

S. G. Mallat. *A wavelet tour of signal processing: the sparse way.* Elsevier /Academic Press, 2009. ISBN 9780123743701.

M. Mani, S. Kurtek, A. Srivastava, and C. Barillot. A comprehensive riemannian framework for the analysis of white matter fiber tracts. In *Proceedings of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1101 –1104, april 2010. doi: 10.1109/ISBI.2010.5490185.

T. W. Manikas and J. T. Cain. Genetic algorithms vs. simulated annealing: A comparison of approaches for solving the circuit partitioning problem. Technical Report 96-101, Department of Electrical Engineering - The University of Pittsburgh, Pittsburgh, PA 15261, May 1996.

R. N. Mantegna and H. E. Stanley. Stochastic process with ultraslow convergence to a gaussian: The truncated lévy flight. *Phys. Rev. Lett.*, 73:2946–2949, Nov 1994. doi: 10.1103/PhysRevLett.73.2946.

A. Matacz. Financial modeling and option theory with the truncated levy process. *International Journal of Theoretical and Applied Finance*, 03(01):143–160, 2000. doi: 10.1142/S0219024900000073.

J. H. McCulloch. Simple consistent estimators of stable distribution parameters. *Communications in Statistics - Simulation and Computation*, 15(4):1109–1136, 1986. doi: 10.1080/03610918608812563.

H. McGurk and J. Macdonald. Hearing lips and seeing voices. *Nature*, 264(5588):746–748, 12 1976.

K. C. McGwire. Cross-validated assessment of geometric accuracy. *Photogrammetric Engineering & Remote Sensing*, 62(10):1179–1187, 1996.

Y. Mekler and Y. J. Kaufman. The effect of earth's atmosphere on contrast reduction·for a nonuniform surface albedo and 'two-halves' field. *Journal of Geophysical Research: Oceans*, 85(C7):4067–4083, July 1980. doi: doi:10.1029/JC085iC07p04067.

A. Messiah. *Quantum Mechanics.* Dover books on physics. Dover Publications, 1999. ISBN 9780486409245. Authors used the French edition: Volumes I & II (ISBN :9782100073610 & 9782100046959), 1995.

K. Miettinen. Application of image quality metrics to problems in remote sensing system design. In *Visual Information Processing XVIII*, volume 5438, pages 150–158. Proc. SPIE, July 2004. doi: 10.1117/12.542136.

G. Miller. A scientist's nightmare: Software problem leads to five retractions. *Science*, 314(5807):1856–1857, 2006. doi: 10.1126/science.314.5807.1856.

H. Mitášová and L. Mitás. Interpolation by regularized spline with tension: I. theory and implementation. *Mathematical Geology*, 25(6):641–655, 1993.

S. Mittnik, T. Doganoglu, and D. Chenyao. Computing the probability density functon of the stable paretian distribution. *Mathematical and Computer Modelling*, 29(10-12):235–240, 1999a. ISSN 0895-7177. doi: 10.1016/S0895-7177(99)00106-5.

S. Mittnik, S. T. rachev, T. Doganoglu, and D. Chenyao. Maximum likelihood estimation of stable paretian models. *Mathematical and Computer Modelling*, 29(10–12):275–293, 1999b. ISSN 0895-7177. doi: 10.1016/S0895-7177(99)00110-7.

H. Moravec and A. Elfes. High resolution maps from wide angle sonar. In *Proceedings of the IEEE International Conference on Robotics and Automation.*, volume 2, pages 116–121, Mar 1985. doi: 10.1109/ROBOT.1985.1087316.

R. D. Morey, J. N. Rouder, J. Verhagen, and E.-J. Wagenmakers. Why hypothesis tests are essential for psychological science: A comment on cumming (2014). *Psychological Science*, 25 (6):1289–1290, 2014. doi: 10.1177/0956797614525969.

J. Morgan. Bayesian estimation for land surface temperature retrieval: the nuisance of emissivities. *IEEE Transactions on Geoscience and Remote Sensing*, 43(6):1279–1288, June 2005. ISSN 0196-2892. doi: 10.1109/TGRS.2005.845637.

J. Morgan. A bayesian estimator for linear calibration error effects in thermal remote sensing. *IEEE Geoscience and Remote Sensing Letters*, 3(1):117–119, 2006. ISSN 1545-598X. doi: 10.1109/LGRS.2005.856755.

R. Morris, X. Descombes, and J. Zerubia. Fully bayesian image segmentation-an engineering perspective. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, volume 3, pages 54–57, Santa Barbara, CA, USA, October 1997.

S. Mukherjee, P. Joshi, S. Mukherjee, A. Ghosh, R. Garg, and A. Mukhopadhyay. Evaluation of vertical accuracy of open source digital elevation model (dem). *International Journal of Applied Earth Observation and Geoinformation*, 21(0):205–217, 2013. ISSN 0303-2434. doi: 10.1016/j.jag.2012.09.004.

R. Müller, M. Lehner, P. Reinartz, and M. Schroeder. Evaluation of spaceborne and airborne line scanner images using a generic ortho-image processor. In *High Resolution Earth Imaging for Geospatial Information, ISPRS Hannover Workshop, Commision I Wg*, volume 5, Mai 2005.

R. Müller, T. Krauss, M. Schneider, and P. Reinartz. Automated georeferencing of optical satellite data with integrated sensor model improvement. *Photogrammetric Engineering and Remote Sensing (PE&RS)*, 78(1):61–74, 2012.

T. Müller. Vergleich von Distanzmetriken beim automatisierten Bildabgleich von hyperspektralen Bilddaten. Bachelor thesis: German Aerospace Center-Julius-Maximilians-Universität - Fakultät für Geographie und Geologie, Lehrstuhl für Fernerkundung, November 2012.

D. E. Myers. Kriging, cokriging, radial basis functions and the role of positive definiteness. *Computers & Mathematics with Applications*, 24(12):139–148, 1992. ISSN 0898-1221. doi: 10.1016/0898-1221(92)90176-I.

D. E. Myers. Spatial interpolation: an overview. *Geoderma*, 62(1-3):17–28, 1994. ISSN 0016-7061. doi: 10.1016/0016-7061(94)90025-6.

E. Næsset. Conditional tau coefficient for assessment of producer's accuracy of classified remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 51(2):91–98, 1996. ISSN 0924-2716. doi: 10.1016/0924-2716(69)00007-4.

P. Nagler, C. Daughtry, and S. Goward. Plant litter and soil reflectance. *Remote Sensing of Environment*, 71(2):207 –215, 2000. ISSN 0034-4257. doi: 10.1016/S0034-4257(99)00082-6.

S. M. Newman, A. M. Larar, W. L. Smith, I. V. Ptashnik, R. L. Jones, M. I. Mead, H. Revercomb, D. C. Tobin, J. K. Taylor, and J. P. Taylor. The joint airborne iasi validation experiment: An evaluation of instrument and algorithms. *Journal of Quantitative Spectroscopy & Radiative Transfer*, 113:1372–1390, 2012. doi: 10.1016/j.jqsrt.2012.02.030.

H. Nguyen and R. Zheng. Binary independent component analysis with or mixtures. *IEEE Transactions on Signal Processing*, 59(7):3168–3181, July 2011. ISSN 1053-587X. doi: 10.1109/TSP.2011.2144975.

NIST - National Institute of Standards and Technology / SEMATECH consortium: NIST/SEMATECH, editor. *e-Handbook of Statistical Methods*. accessible online: http://www.itl.nist.gov/div898/handbook/, version from 30 october 2013 edition, 2013. Last Accessed in Mai 2014.

J. P. Nolan. Numerical calculation of stable densities and distribution functions. *Communications in Statistics. Stochastic Models*, 13(4):759–774, 1997. doi: 10.1080/15326349708807450.

J. P. Nolan. Parameterizations and modes of stable distributions. *Statistics & Probability Letters*, 38(2):187–195, 1998. ISSN 0167-7152. doi: 10.1016/S0167-7152(98)00010-8.

J. P. Nolan. *Stable Distributions - Models for Heavy Tailed Data*. Birkhauser, Boston, 2013. In progress, Chapter 1 online at academic2.american.edu/∼jpnolan.

J. P. Nolan. Financial modeling with heavy-tailed stable distributions. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(1):45–55, 2014. ISSN 1939-0068. doi: 10.1002/wics.1286.

A. Nolte and R. Schrader. A note on the finite time behavior of simulated annealing. *Mathematics of Operations Research*, 25(3):476–484, 2000. doi: 10.1287/moor.25.3.476.12211.

Norsk Elektro Optikk (NEO). Hyspex hyperspectral cameras – an overview. Technical report, Norsk Elektro Optikk (NEO), 2009. Available online: http://www.hyspex.no/hyspex/.

R. Nuzzo. Scientific method: Statistical errors. *Nature*, 506:150–152, February 2014. doi: 10.1038/506150a.

P. Olofsson, G. M. Foody, S. V. Stehman, and C. E. Woodcock. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment*, 129:122–131, 2013. ISSN 0034-4257. doi: 10.1016/j.rse.2012.10.031.

P. Olofsson, G. M. Foody, M. Herold, S. V. Stehman, C. E. Woodcock, and M. A. Wulder. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148(0):42–57, 2014. ISSN 0034-4257. doi: 10.1016/j.rse.2014.02.015.

T. OMP. The OpenMP® API specification for parallel programming. Technical report, The OpenMP Architecture Review Board, http://openmp.org/wp/, January 2012.

M. Ortner, X. Descombes, and J. Zerubia. An adaptive simulated annealing cooling schedule for object detection in images. Technical Report 6336, INRIA - Institut National de Recherche en Informatique et en Automatique, October 2007.

C. Ott, A. Kaldun, P. Raith, K. Meyer, M. Laux, J. Evers, C. H. Keitel, C. H. Greene, and T. Pfeifer. Lorentz meets fano in spectral line shapes: A universal phase and its laser control. *Science*, 340(6133):716–720, 2013. doi: 10.1126/science.1234407.

M. Pagnutti, K. Holekamp, R. Ryan, S. Blonski, R. Sellers, B. Davis, and V. Zanoni. Measurement sets and sites commonly used for characterizations. *International archive of photogrammetry remote sensing and spatial information sciences*, 34(1):159–164, 2002.

P. Pankajakshan, B. Zhang, L. Blanc-Feraud, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia. Blind deconvolution for diffraction-limited fluorescence microscopy. In *Proceedings of the* 5^{th} *IEEE International Symposium on Biomedical Imaging: From Nano to Macro, (ISBI)*, pages 740–743, May 2008. doi: 10.1109/ISBI.2008.4541102.

L. Parra, C. Spence, P. Sajda, A. Ziehe, and K.-R. Müller. Unmixing hyperspectral data. In *in Advances in Neural Information Processing 12 (Proc. NIPS*99*, pages 942–948. MIT Press, 2000.

N. D. Pattengale, M. Alipour, O. R. Bininda-Emonds, B. M. Moret, and A. Stamatakis. How many bootstrap replicates are necessary? *Journal of Computational Biology*, 17(3):337–354, 2010.

J. I. Peltoniemi, S. Kaasalainen, J. Näränen, M. Rautiainen, P. Stenberg, H. Smolander, S. Smolander, and P. Voipio. BRDF measurement of understory vegetation in pine forests: dwarf shrubs, lichen, and moss. *Remote Sensing of Environment*, 94(3):343–354, 2005. ISSN 0034-4257. doi: h10.1016/2004.10.009.

B. T. Phong. Illumination for computer generated pictures. *Commun. ACM*, 18(6):311–317, June 1975. ISSN 0001-0782. doi: 10.1145/360825.360839.

G. Piella and H. Heijmans. A new quality metric for image fusion. In *Proceedings of International Conference on Image Processing, ICIP*, volume 3, pages III–173–6 vol.2, Sept 2003. doi: 10.1109/ICIP.2003.1247209.

J. P. W. Pluim, J. Maintz, and M. Viergever. Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004, Aug 2003. ISSN 0278-0062. doi: 10.1109/TMI.2003.815867.

D. Potts, S. Mackin, J.-P. Muller, and N. Fox. Sensor intercalibration over dome c for the esa globalbedo project. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3):1139–1146, March 2013. ISSN 0196-2892. doi: 10.1109/TGRS.2012.2217749.

E. Povoledo and H. Fountain. Italy orders jail terms for 7 who didn't warn of deadly earthquake. *New-York Times*, October 2012.

M. Powell. Some algorithms for thin plate spline interpolation to functions of two variables. In *International Conference on Advances in Computational Mathematics*, pages 3–18, New Deli, 1993. Department of Applied Mathematics and Theoretical Physics, University of Cambridge.

R. Powell, N. Matzke, C. de Souza Jr., M. Clark, I. Numata, L. Hess, and D. Roberts. Sources of error in accuracy assessment of thematic land-cover maps in the brazilian amazon. *Remote Sensing of Environment*, 90(2):221–234, 2004. ISSN 0034-4257. doi: 10.1016/j.rse.2003.12.007.

F. Pukelsheim and B. Simeone. On the iterative proportional fitting procedure: Structure of accumulation points and l1-error analysis. Technical Report 2009-05, Institut für Mathematik der Universität Augsburg, 2009.

QA4EO. Quality assurance framework for earth observation. http://qa4eo.org/workshops.html, visited in June 2012.

QA4EO task-team. A quality assurance framework for earth observation - principles. Technical report, Committee for Earth-Observation Satellites, 2010.

C. R. N. Rao and J. Chen. Revised post-launch calibration of the visible and near-infrared channels of the advanced very high resolution radiometer (AVHRR) on the NOAA-14 spacecraft. *International Journal of Remote Sensing*, 20(18):3485–3491, 1999. doi: 10.1080/014311699211147.

B. Rasti, J. Sveinsson, M. Ulfarsson, and J. Benediktsson. Hyperspectral image denoising using 3d wavelets. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 1349–1352, July 2012. doi: 10.1109/IGARSS.2012.6351286.

B. Rasti, J. Sveinsson, and M. Ulfarsson. Wavelet-based sparse reduced-rank regression for hyperspectral image restoration. *IEEE Transactions on Geoscience and Remote Sensing*, 52 (10):6688–6698, Oct 2014. ISSN 0196-2892. doi: 10.1109/TGRS.2014.2301415.

B. S. Reddy and B. N. Chatterji. An fft-based technique for translation, rotation, and scale-invariant image registration. *IEEE transactions on image processing*, 5(8):1266–1271, 1996.

R. Reulke and H. Weichelt. SNR evaluation of the rapideye space-borne cameras. *PFG Photogrammetrie, Fernerkundung, Geoinformation*, 2012(1):29–38, 02 2012. doi: 10.1127/1432?8364/2012/0100.

W. H. Richardson. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62(1):55–59, Jan 1972. doi: 10.1364/JOSA.62.000055.

R. Richter. Correction of atmospheric and topographic effects for high spatial resolution satellite imagery. *International Journal of Remote Sensing*, 18(5):1099–1111, 1997. doi: 10.1080/014311697218593.

R. Richter. Bandpass-resampling effects on the retrieval of radiance and surface reflectance. *Applied optics*, 39(27):5001–5005, Sept 2000. doi: 10.1364/AO.39.005001.

R. Richter and D. Schläpfer. Geo-atmospheric processing of airborne imaging spectrometry data. part 2: Atmospheric/topographic correction. *International Journal of Remote Sensing*, 23(13): 2631–2649, 2002. doi: 10.1080/01431160110115834.

R. Richter and D. Schläpfer. Atmospheric / topographic correction for airborne imagery - ATCOR-4 user guide. Technical Report 6.2, DLR-German Aerospace Center and ReSe Applications, 2012. Available online: http://www.rese.ch/download.

R. Richter, M. Bachmann, W. Dorigo, and A. Muller. Influence of the adjacency effect on ground reflectance measurements. *IEEE Geoscience and Remote Sensing Letters*, 3(4):565–569, 2006. ISSN 1545-598X. doi: 10.1109/LGRS.2006.882146.

S. A. Robila. An analysis of spectral metrics for hyperspectral image processing. In *Proceedings of the IEEE international Geoscience and Remote Sensing Symposium (IGARSS)*, volume 5, pages 3233–3236, september 2004. doi: 10.1109/IGARSS.2004.1370390.

M. Rochery, I. H. Jermyn, and J. Zerubia. Higher order active contours. *Int. J. Comput. Vision*, 69(1):27–42, Aug. 2006. ISSN 0920-5691. doi: 10.1007/s11263-006-6851-y.

R. E. Roger and J. F. Arnold. Reliably estimating the noise in aviris hyperspectral images. *International Journal of Remote Sensing*, 17(10):1951–1962, 1996. doi: 10.1080/01431169608948750.

D. Rogge, B. Rivard, J. Zhang, and J. Feng. Iterative spectral unmixing for optimizing per-pixel endmember sets. *IEEE Transactions on Geoscience and Remote Sensing*, 44(12):3725–3736, Dec 2006. ISSN 0196-2892. doi: 10.1109/TGRS.2006.881123.

D. M. Rogge, B. Rivard, J. Zhang, A. Sanchez, J. Harris, and J. Feng. Integration of spatial-spectral information for the improved extraction of endmembers. *Remote Sensing of Environment*, 110(3):287–303, 2007. ISSN 0034-4257. doi: 10.1016/2007.02.019.

K. Rohr, H. Stiehl, R. Sprengel, W. Beil, T. Buzug, J. Weese, and M. Kuhn. Point-based elastic registration of medical image data using approximating thin-plate splines. In K. Höhne and R. Kikinis, editors, *Visualization in Biomedical Computing*, volume 1131 of *Lecture Notes in Computer Science*, pages 297–306. Springer Berlin Heidelberg, 1996. ISBN 978-3-540-61649-8. doi: 10.1007/BFb0046967.

P. Ross and D. Corne. Comparing genetic algorithms, simulated annealing, and stochastic hillclimbing on timetabling problems. In T. Fogarty, editor, *Evolutionary Computing*, volume 993 of *Lecture Notes in Computer Science*, pages 94–102. Springer Berlin Heidelberg, 1995. ISBN 978-3-540-60469-3. doi: 10.1007/3-540-60469-3_28.

D. G. Rossiter. Statistical methods for accuracy assesment of classified thematic maps. Technical report, Department of Earth Systems Analysis International Institute for Geo-information Science & Earth Observation (ITC), Enschede (NL), April 2004.

J. W. Rouse, R. H. Haas, J. A. Schell, and D. W. Deering. Monitoring vegetation systems in the great plains with erts. In *Third ERTS Symposium*, pages 309–317. NASA SP-351, 1973.

J. C. Russ. *The Image Processing Handbook*. CRC Press LLC, 2000 N.W. Corporate Blvd., Boca Raton, FL 33431, USA, 4 edition, 2002. ISBN 0-849-1142-X.

S. Sagar, V. Brando, and M. Sambridge. Noise estimation of remote sensing reflectance using a segmentation approach suitable for optically shallow waters. *IEEE Transactions on Geoscience and Remote Sensing*, 52(12):7504–7512, Dec 2014. ISSN 0196-2892. doi: 10.1109/TGRS.2014.2313129.

R. Saunders, T. Blackmore, B. Candy, P. Francis, and T. Hewison. Monitoring satellite radiance biases using nwp models. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3):1124–1138, March 2013. ISSN 0196-2892. doi: 10.1109/TGRS.2012.2229283.

A. Savitzky and M. J. E. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8):1627–1639, 1964. doi: 10.1021/ac60214a047.

J. Schiewe and C. Kinkeldey. Konzeption und Implementierung einer fallspezifischen Unsicherheitskette im Rahmen fernerkundlicher Auswertungen. *Journal for Photogrammetry, Remote Sensing ans Geoinformation Science*, 6:563–574, December 2013. doi: 10.1127/1432-8364/2013/0200.

D. Schläpfer. PARametric GEocoding, user manual version 3. Technical report, Rese Applications Schläpfer, http://www.rese.ch, 2010.

D. Schläpfer and R. Richter. Geo-atmospheric processing of airborne imaging spectrometry data. part 1: Parametric orthorectification. *International Journal of Remote Sensing*, 23(13):2609–2630, 2002. doi: 10.1080/01431160110115825.

D. Schläpfer, J. Nieke, and K. Itten. Spatial PSF nonuniformity effects in airborne pushbroom imaging spectrometry data. *IEEE Transactions on Geoscience and Remote Sensing*, 45(2): 458–468, Feb 2007. ISSN 0196-2892. doi: 10.1109/TGRS.2006.886182.

E. Schrödinger. *Physique quantique et représentation du monde*, chapter La Situation actuelle en mécanique quantique. Collection Points-Sciences, Le Seuil, 1935. ISBN 2-02-013319-9. Translation in French and edition from 1992.

K. P. Scott, K. J. Thome, and M. R. Brownlee. Evaluation of railroad valley playa for use in vicarious calibration. In *Proceedings of SPIE*, volume 2818, pages 158–166, 1996. doi: 10.1117/12.256090.

W. A. Scott. Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quaterly, Fall*, pages 321–325, 1955.

J. Secker, K. Staenz, R. P. Gauthier, and P. Budkewitsch. Vicarious calibration of airborne hyperspectral sensors in operational environments. *Remote Sensing of Environment*, 76(1): 81–92, 2001. ISSN 0034-4257. doi: 10.1016/S0034-4257(00)00194-2.

M. H. Segall, D. T. Campbell, and M. J. Herskovits. Cultural differences in the perception of geometric illusions. *Science*, 139(3556):769–771, 1963. doi: 10.1126/science.139.3556.769.

M. Sester, H. Hild, and D. Fritsch. Definition of ground-control features for image registration using gis data. *International Archives of Photogrammetry and Remote Sensing*, 32:538–543, 1998.

Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *ACM Trans. Graph.*, 27(3):73:1–73:10, Aug. 2008. ISSN 0730-0301. doi: 10.1145/1360612.1360672.

G. Shao and J. Wu. On the accuracy of landscape pattern analysis using remote sensing data. *Landscape Ecology*, 23(5):505–511, 2008. ISSN 0921-2973. doi: 10.1007/s10980-008-9215-x.

S. A. Shappel and D. A. Wiegmann. The human factor analysis and classification system-HFACS. Technical report, Office for Aviation Medicine, Federal Aviation Administration, US Department of Transportation, FAA Civil Aeromedical Institute, Oklahoma City, OK 73125, February 2000.

K. Sharifi and A. Leon-Garcia. Estimation of shape parameter for generalized gaussian distributions in subband decompositions of video. *IEEE Transactions on Circuits and Systems for Video Technology*, 5(1):52–56, 1995. ISSN 1051-8215. doi: 10.1109/76.350779.

L. Shi and J. J. Bates. Three decades of intersatellite-calibrated high-resolution infrared radiation sounder upper tropospheric water vapor. *Journal of Geophysical Research: Atmospheres*, 116 (D4), 2011. ISSN 2156-2202. doi: 10.1029/2010JD014847.

A. Shnayderman, A. Gusev, and A. Eskicioglu. An svd-based grayscale image quality measure for local and global assessment. *IEEE Transactions on Image Processing*, 15(2):422–429, Feb 2006. ISSN 1057-7149. doi: 10.1109/TIP.2005.860605.

B. W. Silverman. Spline smoothing: The equivalent variable kernel method. *The Annals of Statistics*, 12(3):898–916, 09 1984. doi: 10.1214/aos/1176346710.

J. Sim and C. C. Wright. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy*, 85(3):257–268, 2005.

K. Simonson, S. Drescher, and F. Tanner. A statistics-based approach to binary image registration with uncertainty analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(1):112–125, Jan 2007. ISSN 0162-8828. doi: 10.1109/TPAMI.2007.250603.

D. A. Sims and J. A. Gamon. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, 81(2–3):337–354, 2002. ISSN 0034-4257. doi: 10.1016/S0034-4257(02)00010-X.

D. Six, M. Fily, S. Alvain, P. Henry, and J.-P. Benoist. Surface characterisation of the dome concordia area (antarctica) as a potential satellite calibration site, using spot-4 vegetation instrument. *Remote Sensing of Environment*, 89(1):83–94, 2004. ISSN 0034-4257. doi: 10.1016/j.rse.2003.10.006.

P. C. Smits, S. G. Dellepiane, and R. A. Schowengerdt. Quality assessment of image classification algorithms for land-cover mapping: A review and a proposal for a cost-based approach. *International Journal of Remote Sensing*, 20(8):1461–1486, 1999. doi: 10.1080/014311699212560.

I. Solheim, O. Engelsen, B. Hosgood, and G. Andreoli. Measurement and modeling of the spectral and directional reflection properties of lichen and moss canopies. *Remote Sensing of Environment*, 72(1):78–94, 2000. ISSN 0034-4257. doi: 10.1016/S0034-4257(99)00093-0.

J. Soukkamaki. The first generation of 2d hyperspectral camera is airborne. Technical report, Rikola Ldt., Kaitoväylä 1, FIN-90570 Oulu, Finland, 2012.

Specim. AISADual datasheet. Technical report, Spectral Imaging Ltd., POB 110, Teknologiantie 18 A, FIN–90571 Oulu, Finland, 2012. version-1.2012.

A. Srivastava, S. H. Joshi, W. Mio, and X. Liu. Statistical shape analysis: Clustering, learning, and testing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):590 –602, april 2005. ISSN 0162-8828. doi: 10.1109/TPAMI.2005.86.

A. Srivastava, E. Klassen, S. Joshi, and I. Jermyn. Shape analysis of elastic curves in euclidean spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(7):1415–1428, July 2011. ISSN 0162-8828. doi: 10.1109/TPAMI.2010.184.

J. L. Starck, E. Pantin, and F. Murtagh. Deconvolution in astronomy: A review. *PASP: Publications of the Astronomical Society of the Pacific*, pages 1051–1069, October 2002.

S. Stehman. Estimating the kappa coefficient and its variance under stratified random sampling. *Photogrammetric Engineering and Remote Sensing*, 62(4):401–407, 1996.

S. V. Stehman. Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 62(1):77–89, 1997. ISSN 0034-4257. doi: 10.1016/S0034-4257(97)00083-7.

S. V. Stehman. A critical evaluation of the normalized error matrix in map accurcy assessment. *Photogrammetric Engineering and Remote Sensing*, 70(6):743–751, 2004.

S. V. Stehman. Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing*, 30(20):5243–5272, 2009. doi: 10.1080/01431160903131000.

S. V. Stehman and R. L. Czaplewski. Design and analysis for thematic map accuracy assessment: Fundamental principles. *Remote Sensing of Environment*, 64(3):331–344, 1998. ISSN 0034-4257. doi: 10.1016/S0034-4257(98)00010-8.

D. Steinvorth. Finding swimming pools with google earth: Greek government hauls in billions in back taxes. online article published by Spiegel Online International, August 2010. Accessible online: http://www.spiegel.de/international/europe/, last accessed in June 2014.

S. Stevens. On the theory of scales of measurement. *Science*, 103(2684):677–680, June 1946. doi: 10.1126/science.103.2684.677.

G. Stockman, S. Kopstein, and S. Benett. Matching images to models for registration and object detection via clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-4(3):229–241, May 1982. ISSN 0162-8828. doi: 10.1109/TPAMI.1982.4767240.

T. Stone, W. Rossow, J. Ferrier, and L. Hinkelman. Evaluation of isccp multisatellite radiance calibration for geostationary imager visible channels using the moon. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3):1255–1266, March 2013. ISSN 0196-2892. doi: 10.1109/TGRS.2012.2237520.

G. Strang and V. Strela. Short wavelets and matrix dilation equations. *IEEE Transactions on Signal Processing*, 43(1):108–115, Jan 1995. ISSN 1053-587X. doi: 10.1109/78.365291.

T. Stuffler, H. Kaufmann, S. Hofer, K.-P. Forster, G. Schreier, A. Müller, A. Eckardt, H. Bach, B. Penne, U. Benz, and R. Haydn. The EnMAP hyperspectral imager- an advanced optical payload for future applications in earth observation programmes. *Acta Astronautica*, 61:115–120, 2007. doi: 10.1016/j.actaastro.2007.01.033.

A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics (SIAM), 2005. ISBN 0-89871-572-5.

T. Teršič, U. Abdybachaev, F. Blanchard, G. Cheban, C. Ehrler, W. E. Falck, C. Fischer, S. R. Grebby, H. Hejny, C. J. Jordan, G. Kerr, E. Kylychbaev, E. Mambetaliyev, V. Mardhel, F. M. McEvoy, O. Rouzeau, R. Šajn, S. Šolar, D. Wittmer, G. Žibret, and S. Chevrel. Indicators and earth observation products for the assessment of the extractive industry environmental and societal impacts - material and information for the eo-miners workshop at the makmal gold deposit demonstration site., 2013a. doi: 10.5474/eo-miners.2013.03-eng. EO-MINERS project,http://www.eo-miners.eu/.

T. Teršič, F. Andriamasinoro, E. Ben-Dor, F. Blanchard, A. Bourguignon, H. Coetzee, C. Ehrler, W. Falck, S. Ferriday, C. Fischer, S. R. Grebby, B. Hanise, H. Hejny, C. J. Jordan, G. Kerr, M. Kwata, I. Livno, V. Mardhel, F. M. McEvoy, J. Mišurec, G. Nosteco, B. Palumbo-Roe, D. Rogge, O. Rouzeau, S. V. šolar, D. Van Tonder, D. Wittmer, G. Žibret, L. Žibret, and S. Chevrel. Indicators and earth observation products for the assessment of the extractive industry environmental and societal impacts - material and information for the eo-miners workshop at the emalahleni coalfield demonstration site, 2013b. doi: 10.5474/eo-miners.2013.02.

. TET: Technologieerprobungsträger. Tet-1 information page, 2014. http://www.dlr.de/rd/desktopdefault.aspx/tabid-2274//3396_read-5085/Technologieerprobungstr%C3%A4ger.

S. F. B. Tett, P. A. Stott, M. R. Allen, W. J. Ingram, and J. F. B. Mitchell. Causes of twentieth-century temperature change near the earth's surface. *Nature*, 399(6736):569–572, 1999. ISSN 0028-0836. doi: 10.1038/21164.

P. S. Thenkabail, R. B. Smith, and E. De Pauw. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, 71(2): 158–182, 2000. ISSN 0034-4257. doi: 10.1016/S0034-4257(99)00067-X.

H. C. S. Thom. A note on the gamma distribution. *Monthly Weather Review*, 86(4):117–122, April 1958. ISSN 0027-0644. doi: 10.1175/15200493.1958.0860117.

K. Thome, J. McCorkel, and J. Czapla-Myers. In-situ transfer standard and coincident-view intercomparisons for sensor cross-calibration. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3):1088–1097, March 2013. ISSN 0196-2892. doi: 10.1109/TGRS.2013.2243841.

I. D. Thompson, S. C. Maher, D. P. Rouillard, J. M. Fryxell, and J. A. Baker. Accuracy of forest inventory mapping: Some implications for boreal forest management. *Forest Ecology and Management*, 252(1-3):208–221, 2007. ISSN 0378-1127. doi: 10.1016/j.foreco.2007.06.033.

D. C. Tobin, H. E. Revercomb, R. O. Knuteson, F. A. Best, W. L. Smith, N. N. Ciganovich, R. G. Dedecker, S. Dutcher, S. D. Ellington, R. K. Garcia, H. B. Howell, D. D. LaPorte, S. A. Mango, T. S. Pagano, J. K. Taylor, P. van Delst, K. H. Vinson, and M. W. Werner. Radiometric and spectral validation of atmospheric infrared sounder observations with the aircraft-based scanning high-resolution interferometer sounder. *Journal of Geophysical Research: Atmospheres*, 111(D9), 2006. ISSN 2156-2202. doi: 10.1029/2005JD006094.

V. Černý. Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of Optimization Theory and Applications*, 45(1):41–51, 1985. ISSN 0022-3239. doi: 10.1007/BF00940812.

P. Čížek, R. Weron, and W. Härdle. *Statistical tools for Finance and Insurance*, chapter 1: Stable Distributions, pages 21–44. Springer, 2005. ISBN 978-3-540-22189.

J. S. Uebersax. Diversity of decision-making models and the measurement of interrater agreement. *Psychological Bulletin*, 1987.

S. L. Ustin, editor. *Remote Sensing for Natural Resource Management and Environmental Monitoring*, volume 4. John Wiley & Sons, Inc., University of California, Davis, third edition, 2004. ISBN 0-471-31793-4. Published in cooperation with the American Society for Photogrammetry and Remote Sensing (ASPRS).

G. Žibret, S. Adar, E. Ben-Dor, A. Bourguignon, C. Ehrler, E. Falck, C. Fischer, S. R. Grebby, H. Hejny, J. Jelenek, C. J. Jordan, G. Kerr, V. Kopačkova, I. Livne, V. Mardhel, F. M. McEvoy, J. Mišurec, G. Notesco, B. Palumbo-Roe, D. Rogge, S. šolar, T. Terši, D. Wittmer, and S. Chevrel. Indicators and earth observation products for the assessment of the extractive industry environmental and societal impacts-material and information for the eo miners workshop at the sokolovská uhelná demonstration site., 2013. doi: 10.5474/eo-miners.2013.01. EO-MINERS Project: http://www.eo-miners.eu/.

E. Vermote, R. Santer, P. Y. Deschamps, and M. Herman. In-flight calibration of large field of view sensors at short wavelengths using rayleigh scattering. *International Journal of Remote Sensing*, 13(18):3409–3429, 1992. doi: 10.1080/01431169208904131.

A. J. Viera and J. M. Garrett. Understanding interobserver agreement: The kappa statistic. *Family Medicine*, 37(5):360–363, May 2005.

L. Vincent. Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. *IEEE Transactions on Image Processing*, 2(2):176–201, Apr 1993. ISSN 1057-7149. doi: 10.1109/83.217222.

P. Viola and W. Wells. Alignment by maximization of mutual information. In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pages 16–23, Jun 1995. doi: 10.1109/ICCV.1995.466930.

R. von der Heydt, E. Peterhans, and G. Baumgartner. Illusory contours and cortical neuron responses. *Science*, 224(4654):1260–1262, 1984. doi: 10.1126/science.6539501.

G. Wahba. *Spline Models for Observational Data*. SIAM - Society for Industrial and Applied Mathematics, 1990. doi: 10.1137/1.9781611970128.

E. H. Walker. A mathematical theory of optical illusions and figural aftereffects. *Perception & Psychophysics*, 13(3):467–486, 1973. ISSN 0031-5117. doi: 10.3758/BF03205804.

Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004. ISSN 1057-7149. doi: 10.1109/TIP.2003.819861.

C. Wei and R. S. Blum. Theoretical analysis of correlation-based quality measures for weighted averaging image fusion. *Information Fusion*, 11(4):301–310, 2010. ISSN 1566-2535. doi: 10.1016/j.inffus.2009.10.006.

S. Weide, M. Bachmann, P. Gege, C. Schwarz, S. Holzwarth, and A. Müller. OpAiRS - optical airborne remote sensing and calibration facility. In *Proceedings of Hyperspectral Workshop 2010*, Frascati, Italy, March 2010. ESA Communications SP-683, Estec, Noordwjik, The Netherlands. ISBN 978-92-9221-247-6.

T. Weise. Global optimization algorithms - theory and application. Technical report, Nature Inspired Computation and Applications Laboratory (NICAL) of the University of Science and Technology of China (USTC), June 2009. Available online: http://www.it-weise.de/.

R. Wiemker, K. Rohr, R. Sprengel, and H. S. Stiehl. Application of elastic registration to imagery from airborne scanners. In *In Proceedings of the XVIII. Congress of the International Society for Photogrammetry and Remote Sensing ISPRS 1996, Vienna, volume XXXI part B4 of International Archives of Photogrammetry and Remote Sensing*, pages 949–954, 1996.

G. Wilkinson. Results and implications of a study of fifteen years of satellite image classification experiments. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):433–440, March 2005. ISSN 0196-2892. doi: 10.1109/TGRS.2004.837325.

D. P. Wipf and H. Zhang. Revisiting bayesian blind deconvolution. *CoRR*, abs/1305.2362, 2013.

N. Wongpakaran, T. Wongpakaran, D. Wedding, and K. L. Gwet. A comparison of Cohen's kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Medical Research Methodology*, 13(61):1–7, 2013.

O. Yilmaz and S. Rickard. Blind separation of speech mixtures via time-frequency masking. *IEEE Transactions on Signal Processing*, 52(7):1830–1847, July 2004. ISSN 1053-587X. doi: 10.1109/TSP.2004.828896.

Y.-L. You and M. Kaveh. A regularization approach to joint blur identification and image restoration. *IEEE Transactions on Image Processing*, 5(3):416–428, Mar 1996. ISSN 1057-7149. doi: 10.1109/83.491316.

P. Zarco-Tejada, A. Berjón, R. López-Lozano, J. Miller, P. Martín, V. Cachorro, M. González, and A. de Frutos. Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sensing of Environment*, 99(3):271–287, 2005. ISSN 0034-4257. doi: 10.1016/j.rse.2005.09.002.

V. Zarzoso, P. Comon, and R. Phlypo. A contrast function for independent component analysis without permutation ambiguity. *IEEE Transactions on Neural Networks*, 21(5):863–868, May 2010. ISSN 1045-9227. doi: 10.1109/TNN.2010.2045128.

J. Zhang and G. M. Foody. A fuzzy classification of sub-urban land cover from remotely sensed imagery. *International Journal of Remote Sensing*, 19(14):2721–2738, 1998. doi: 10.1080/014311698214479.

T. Zhang and W. Luo. Image quality assessment using author topic model. In *International Conference on Information Technology and Applications (ITA)*, pages 63–66, Nov 2013. doi: 10.1109/ITA.2013.21.

Q. Zheng and R. Chellappa. A computational vision approach to image registration. In *Pattern Recognition, 1992. Vol.I. Conference A: Computer Vision and Applications, Proceedings., 11th IAPR International Conference on*, pages 193–197, Aug 1992. doi: 10.1109/ICPR.1992.201539.

M. Zibulevsky and B. A. Pearlmutter. Blind source separation by sparse decomposition in a signal dictionary. *Neural Comput.*, 13(4):863–882, Apr. 2001. ISSN 0899-7667. doi: 10.1162/089976601300014385.

B. Zitová and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003. ISSN 0262-8856. doi: 10.1016/S0262-8856(03)00137-9.

Appendix A

Notations

> Mathematics is a game played according to certain simple rules with meaningless marks on paper.
>
> David Hilbert - Mathematician

The following notations are given to facilitate the understanding of readers not familiar with mathematics. The goal is therefore not to provide sound mathematical definitions of these terms, but only to give an insight into the notations and involved concepts.

A.1 Notations for Sets

Symbol	Explanation	Example
$\mathbb{N}$	Set of positive integers	e.g. 0, 1, 2,...
$\mathbb{Z}$	Set of all integers	e.g. -2, -1, 0, 1, 2
$\mathbb{R}$	Set of real numbers	e.g. 0.123, 1/3, π
$\mathbb{C}$	Set of complex numbers	e.g. $i = \sqrt{-1}$
$\mathbb{L}^p(\mathbb{S})$	Set of functions whose power of p is integrable over $\mathbb{S}$	e.g. $x \to \exp -x^2 \in \mathbb{L}^2(\mathbb{R})$
$\mathbb{S}^\star$	All elements from $\mathbb{S}$ except 0	e.g. $\mathbb{N}^\star = \{1, 2, 3, \ldots\}$
$\mathbb{S}^+$	All $\mathbb{S}$ except its negative elements	e.g. $\mathbb{Z}^+ = \mathbb{N}$
$\mathbb{S}^-$	All $\mathbb{S}$ except its positive elements	e.g. $\mathbb{Z}^- = \{0, -1, -2, \ldots\}$
$\mathbb{S}\backslash\mathbb{T}$	Elements from $\mathbb{S}$ but not in $\mathbb{T}$	e.g. $\mathbb{Z}\backslash\mathbb{N} = \{-1, -2, \ldots\}$
$[\![a; b]\!]$	Integers between a and b	$[\![0, 2]\!] = \{0, 1, 2\}$
$[a; b]$	Real numbers between a and b	$[3, 4] = \{3, \ldots \pi, \ldots 4\}$

Table A.1: Notations associated with sets.

A.2 Notations for Numbers

Symbol	Explanation	Illustration
i	Imaginary unit	$i^2 = -1$
e	Exponential number	$e = \sum_{n \in \mathbb{N}} \frac{1}{n!} \approx 2.7182\ldots$
π	Pi, half perimeter of a circle of radius 1	$\pi \approx 3.1415\ldots$
$\lvert x \rvert$	Absolute Value or Modulus of x	$\lvert -1 \rvert = 1$, $\lvert 3 + 4i \rvert = 5$
$arg(c)$	Phase of c	$arg(4i) = \frac{\pi}{2}$
$\overline{c}$	Complex conjugate of c	$\overline{3 + 4i} = 3 - 4i$
$\mathcal{R}(c)$	Real part of c	$\mathcal{R}(3 + 4i) = 3$
$\mathcal{I}(c)$	Imaginary part of c	$\mathcal{I}(3 + 4i) = 4$
$\lfloor X \rfloor$	Floor of X	$\lfloor 4.5 \rfloor = 4$
$\lceil X \rceil$	Ceiling of X	$\lceil 4.5 \rceil = 5$

Table A.2: Notations associated with numbers.

A.3 Miscellaneous Notations

Symbol	Explanation	Illustration
X^t	Transposition of matrix X	e.g. $(1, 2)^t = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$
∇	Gradient of a function	e.g. $\nabla f(x, y) = \begin{pmatrix} \frac{\partial f}{\partial x}(x, y) \\ \frac{\partial f}{\partial y}(x, y) \end{pmatrix}$
!	Factorial operator	$n! = 1 \times 2 \times \ldots \times (n-1) \times n$
$P(X\|Y)$	The probability to observe X given Y	$P(\text{long-hair}\|\text{women}) > P(\text{long-hair}\|\text{man})$
$\hat{f}$ and $\check{f}$	Fourier and inverse Fourier transform of f	see appendix D
$f \star g$	Convolution of f and g	$(f \star g)(x) = \int_{\infty}^{\infty} f(t)g(x-t)dt$

Table A.3: Miscellaneous notations.

A.4 Special Functions

This section reviews some specific functions used within this work. They are presented by alphabetic order.

A.4.1 Beta Function

The expression of the beta function, also called Euler integral, is given in equation A.4.1.

$$B : \left\{ \begin{array}{ccc} \mathbb{R}^{+\star} \times \mathbb{R}^{+\star} & \rightarrow & \mathbb{R}^{+} \\ (x,y) & \rightarrow & \int_0^1 t^{x-1}\left(1-t\right)^{y-1} dt \end{array} \right. \tag{A.4.1}$$

A.4.2 Complementary Error Function

The complementary error function is defined in equation A.4.2.

$$\text{erfc} : \left\{ \begin{array}{ccc} \mathbb{R} & \rightarrow & [0;2] \\ x & \rightarrow & \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-t^2} dt \end{array} \right. \tag{A.4.2}$$

A.4.3 Di-Gamma Function

The di-gamma function is commonly denoted by Ψ, it correspond to the logarithmic derivative of the gamma function (see section A.4.4).

$$\Psi : \left\{ \begin{array}{ccc} \mathbb{R} \backslash \mathbb{Z}^{-} & \rightarrow & \mathbb{R} \\ x & \rightarrow & \frac{1}{\Gamma(x)} \frac{d\Gamma}{dx}(x) \end{array} \right. \tag{A.4.3}$$

Beside many interesting properties, it can be shown that for positive x, $\Psi(x) \in [\log(x-1); \log(x)]$, thus yielding the approximation equation A.4.4.

$$\Psi(x) \approx \ln(x - 0.5) \tag{A.4.4}$$

A.4.4 Gamma Function

The gamma function is defined as an extension of the factorial: $\Gamma(n) = (n-1)!$. Although it can be extended to complex and negative real numbers, its expression is only given for positive real numbers.

$$\Gamma : \left\{ \begin{array}{ccc} \mathbb{R}^{+\star} & \rightarrow & \mathbb{R} \\ x & \rightarrow & \int_0^{\infty} t^{x-1} e^{-t} dt \end{array} \right. \tag{A.4.5}$$

A.4.5 Lagrange Polynomial

The Lagrange polynomials are a family of interpolating polynomials. Given a set of N control points, $\{(x_k, y_k)\}_k$, the Lagrange polynomial is the polynomial of least degree (i.e. $N-1$) being equal to y_k at each x_k:

$$L : \left\{ \begin{array}{ccc} \mathbb{R} & \rightarrow & \mathbb{R} \\ x & \rightarrow & \sum\limits_{k=1}^{N} y_k \left(\prod\limits_{j=1}^{k-1} \frac{x - x_j}{x_k - x_j} \right) \left(\prod\limits_{j=k+1,}^{N} \frac{x - x_j}{x_k - x_j} \right) \end{array} \right. \tag{A.4.6}$$

A.4.6 Kronecker Function

This function merely indicates whether two elements are equals or not, thus its alternate name 'indicator function'. It is given in equation A.4.7.

$$\delta : \begin{cases} \mathbb{R} \times \mathbb{R} & \rightarrow & \{0,1\} \\ (x,y) & \rightarrow & \begin{cases} 0 & \text{if} & x \neq y \\ 1 & \text{if} & x = y \end{cases} \end{cases} \tag{A.4.7}$$

It is usually applied to a subset of $\mathbb{R}$, typically $\mathbb{N}$. In order to simplify notations, the symbol δ_{xy} is used instead of $\delta(x,y)$.

A.4.7 Sign Function

The sign function simply accounts for the sign of a real number. Its formal definition is given by equation A.4.8.

$$sgn : \begin{cases} \mathbb{R} & \rightarrow & \{-1,0,1\} \\ & & \\ x & \rightarrow & \begin{cases} -1 & \text{if} & x < 0 \\ 0 & \text{if} & x = 0 \\ 1 & \text{if} & x > 0 \end{cases} \end{cases} \tag{A.4.8}$$

A.4.8 Tri-Gamma Function

The tri-gamma function corresponds to the second logarithmic derivative of the gamma function:

$$\psi'(x) = \frac{\partial^2}{\partial x^2} \log(\Gamma)(x) \tag{A.4.9}$$

It can be defined and extended by means of series (see equation A.4.10).

$$\psi' : \begin{cases} \mathbb{C} & \rightarrow & \mathbb{R} \\ x & \rightarrow & \sum_{k=0}^{\infty} \frac{1}{(x+n)^2} \end{cases} \tag{A.4.10}$$

A.5 Physical Constants

Notation	Name	Value	Units (ISU)
c_0	Speed of light in vacuum	$299792458 \approx 3 \cdot 10^8$	$m \cdot s^{-1}$
n_0	Refraction index of the air	$\approx 1.0003 \approx 1$	N/A
k_B	Boltzmann's constant	$1.38065052 \cdot 10^{-23}$	$J \cdot K^{-1}$
h	Planck's constant	$6.626069 \cdot 10^{-34}$	Js
R	Gas constant	8.314462	$J \cdot K^{-1} mol^{-1}$

Table A.4: Summary of physical constants used in this work. Please note that their 'values' might depend on several parameters: only typical ranges are provided here.

A.5.1 Note on Refraction Index

The refraction index of an optical medium is a dimensionless number. It is defined as the ratio between the speed of light in the medium, c, and the speed of light in vacuum c_0 (see equation A.5.1). It depends of the radiation wavelength. The refractive index of air and atmosphere is commonly approximated by 1, thus 'granting' a similar behaviour as in a vacuum.

$$n(\lambda) = \frac{c_0}{c} \tag{A.5.1}$$

A.5.2 Ideal Gas Law

The ideal gas law links the pressure (P) and volume V of a gas to its absolute temperature (T) and its amount of substance (n) by means of the gas constant R: see equation A.5.2.

$$PV = nRT \tag{A.5.2}$$

Appendix B

Detailed results

The Beaver brought paper, portfolio, pens,
And ink in unfailing supplies:
While strange creepy creatures came out of their dens,
And watched them with wondering eyes.

So engrossed was the Butcher, he heeded them not,
As he wrote with a pen in each hand,
And explained all the while in a popular style
Which the Beaver could well understand.

"Taking Three as the subject to reason about -
A convenient number to state -
We add Seven, and Ten, and then multiply out
By One Thousand diminished by Eight.

"The result we proceed to divide, as you see,
By Nine Hundred and Ninety Two:
Then subtract Seventeen, and the answer must be
Exactly and perfectly true.

"The method employed I would gladly explain,
While I have it so clear in my head,
If I had but the time and you had but the brain-
But much yet remains to be said."

Lewis Caroll - Mathematician

B.1 Sokolov 2010 Classification

B.1.1 Cross-Confusion Matrixes

Classes from Flight-line 2 \ Classes from Flight-line 1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	2208	0	0	0	1	0	4	1	184	0	38	29	1	46	0	0	0	69	224	183	0	10	16	202	0	0	0	2	2	6	0	0	68	1	0	0	2	0	0	6	32	0	1	13	0	0	0	0	0	9	0	0	60	0	0	32	0	11	0	0
3	0	0	237	31	4	0	0	3	0	0	0	0	0	0	143	13	0	0	0	0	0	0	1	0	2	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0
4	0	0	14	114	0	0	0	0	0	0	0	0	0	0	51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0
5	0	0	2	1	327	0	0	0	14	0	0	0	0	0	115	2	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	5	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
7	0	0	0	0	0	0	192	51	0	2	0	1	65	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	14	0	1	3	0	75	2336	79	532	0	63	321	5	178	0	0	0	2	0	0	0	0	0	3	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	170	0	0	0	0	0	0	0	0
9	0	1	0	0	199	0	2	19	430	0	0	0	8	0	175	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0
10	0	396	0	0	0	0	2	440	0	10674	0	21	103	2	652	0	0	0	2	35	53	0	2	0	735	0	0	0	0	0	0	0	0	70	3	5	0	5	0	0	1	99	2	0	7	0	0	0	0	0	9	0	0	971	0	4	10	0	22	3	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	38	0	0	0	0	0	64	0	18	2	484	46	0	2	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
13	0	53	0	0	1	4	7	91	7	71	0	21	1263	0	263	0	0	0	0	1	9	0	1	0	37	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	45	0	1	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	97	41	51	163	3	10	60	31	423	0	2	113	1	7452	473	3	0	1	101	26	1	802	0	1937	0	0	0	0	0	0	0	0	3	38	3	0	1	0	0	0	1	1	0	1	0	0	0	0	0	0	1	1	327	0	5	165	0	16	0	0
16	0	2	9	1	3	0	0	0	0	0	0	0	1	0	51	87	1	0	0	18	0	0	124	0	17	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	654	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	7880	1153	154	1	1	57	36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	1489	0	0	0	0	0	0	0	23	0	1	2	0	34	3	0	0	4402	23526	499	41	307	66	793	0	0	0	0	0	0	0	0	5	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	1	16	0	0	0	0	0	0
21	0	433	0	0	0	0	0	0	0	61	0	3	3	0	18	0	0	0	2156	1717	3263	0	5	14	706	0	0	0	0	0	0	0	0	2	3	2	0	0	0	0	0	1	1	2	0	0	0	0	0	0	1	0	0	18	0	1	12	0	3	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	31	0	0	1	0	0	0	0	2	0	0	1	0	38	21	1	0	2	974	4	14	974	2	118	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0
24	0	284	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	54	23	1	0	0	59	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	629	0	0	0	0	1	0	0	390	0	1	23	0	728	21	0	0	107	2565	971	1	1075	4	7953	0	0	0	0	0	0	0	0	5	57	5	0	1	0	0	0	10	1	0	2	0	0	0	1	0	3	0	1	321	0	10	120	0	22	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	3	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	18	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0	1	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	63	0	0	0	0	0	2	0	105	0	2	2	1	2	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	324	0	0	0	0	0	0	1	11	0	0	1	0	0	0	1	0	7	0	0	37	0	0	0	0	0	0	0	0
35	0	5	3	0	0	0	0	0	0	0	0	0	1	0	18	3	0	0	0	2	2	0	3	0	37	0	0	0	0	0	0	0	0	0	69	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	1	0	0	0	0
36	0	6	0	0	0	0	0	0	0	2	0	0	2	0	5	0	0	0	0	2	2	0	0	0	3	0	0	0	0	0	0	0	0	1	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	2	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	0	17	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
42	0	17	0	0	0	0	0	0	0	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	2	0	0	0	0	0	0	0	0	26	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	1	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0
44	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	0	5	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	0	0	9	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	2	14	0	0	0	0	0	0	0	0
48	0	2	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
51	0	8	0	0	0	0	0	0	0	20	0	0	1	0	1	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	10	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	4	0	1	0	0	6	0	0
52	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
54	0	72	5	7	0	0	0	48	0	291	0	0	5	0	330	0	0	0	1	8	10	0	1	0	318	0	0	0	0	0	0	0	0	5	12	0	0	2	0	0	0	13	1	0	4	0	0	0	1	0	0	0	1998	0	0	2	0	0	9	0
55	0	0	0	0	0	0	0	4	0	10	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	5	0	0	0	2	0	0	0	0	1	0	12	32	0	0	0	0	0	0	0
56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	21	0	0	7	0	0	0
57	0	29	0	0	0	0	0	0	0	7	0	1	10	0	56	3	0	0	0	29	23	0	17	0	130	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	223	0	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
59	0	26	0	0	0	0	0	0	0	21	0	0	2	0	20	0	0	0	2	3	1	0	0	1	22	0	0	0	0	0	0	0	0	0	2	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	4	6	0	11	0	0	89	0	0	0
60	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0
61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table B.1: Sokolov 2010, confusion matrix between flight-lines 1 and 2. Total matches: 112235, $\kappa = 0.59$, $\hat{\sigma}_\kappa = 0.0017$. See table 6.2 for classes index definitions.

Classes from Flight-line 3 \ Classes from Flight-line 2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	2968	1	2	0	2	0	1	0	336	0	58	7	3	24	0	0	0	65	160	70	0	5	8	125	0	0	0	0	0	0	7	0	22	1	0	0	4	0	0	0	5	0	0	1	0	0	0	0	0	2	0	1	43	0	1	7	0	6	0	0
3	0	6	968	101	37	0	1	26	0	16	0	2	0	0	279	155	0	0	0	0	0	0	16	0	4	0	0	0	0	0	0	0	0	0	76	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	78	0	1	0	0	0	0	0
4	0	11	37	111	0	0	0	15	0	10	0	0	0	0	68	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	0	0
5	0	0	57	1	20	0	0	0	0	0	0	0	0	0	98	28	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	2	0	0	0	45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0
7	0	0	0	0	0	0	1	5	0	1	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	3	0	0	0	2	0	0	0
8	0	14	2	0	0	0	1	639	2	344	1	7	14	2	57	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	53	0	0	0	2	0	0	0
9	0	0	0	0	0	0	0	1	5	0	0	0	2	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	607	2	0	0	0	0	249	0	43715	16	178	41	43	685	0	0	0	2	35	34	0	3	0	728	0	0	0	0	0	0	2	0	46	17	0	0	6	0	0	0	4	2	0	8	0	0	0	0	0	39	0	0	711	1	12	2	0	63	0	0
11	0	13	0	0	0	0	0	0	0	39	7	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	77	0	0	0	0	0	5	0	405	0	274	0	3	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
13	0	18	0	0	0	0	0	0	0	24	0	0	28	0	42	0	0	0	0	2	3	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	9	0	0	2	0	0	3	0	3	0	0
14	0	4	0	0	0	0	0	0	0	16	0	2	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
15	0	88	102	12	30	0	0	59	0	277	0	2	11	3	3323	544	5	0	1	96	10	1	533	0	1348	0	0	0	0	0	0	0	0	0	133	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	11	158	1	3	51	0	10	0	0
16	0	3	34	0	1	0	0	0	0	0	0	0	0	0	31	234	1	0	0	9	1	0	118	0	16	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	613	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	8493	1300	59	2	2	103	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
20	0	1294	0	0	0	0	0	0	0	7	0	0	1	0	14	6	0	0	3995	21602	131	119	221	136	474	0	0	0	0	0	0	0	0	0	4	1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	7	0	0	1	0	1	0	0
21	0	455	0	0	0	0	0	0	0	27	0	0	0	0	3	1	0	0	863	1304	677	0	5	42	376	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	1	0	2	0	0
22	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	18	8	0	0	0	0	0	0	0	0	0	0	0	10	40	3	0	9	1548	0	15	988	0	76	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3	0	0	0	0
24	0	119	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	39	22	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	562	3	0	0	0	0	0	0	205	0	0	2	0	273	59	0	0	135	2687	369	0	939	5	4493	0	0	0	0	0	0	0	0	0	30	2	0	17	0	0	0	5	0	0	1	0	0	0	0	0	3	0	1	235	0	5	25	0	33	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	6	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	23	0	2	0	0	0	0	0	63	0	0	0	1	3	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	86	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	12	0	0	0	0	0	0	0
35	0	1	12	0	5	0	0	0	0	2	0	0	0	0	36	25	0	0	0	1	0	0	8	0	25	0	0	0	0	0	0	0	0	0	94	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	2	0	0	0	0	0
36	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
42	0	14	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	2	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0
43	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	0	5	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	4	0	0	31	0	0	0	0	0	0	0	0	26	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
51	0	15	0	0	0	0	0	0	0	24	0	0	6	0	6	0	0	0	0	1	2	0	0	0	7	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	47	1	0	4	0	0	0	6	0	0	0
52	0	2	0	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
54	0	153	14	10	0	0	2	68	0	251	0	5	8	0	522	0	0	0	1	33	13	0	7	0	566	0	0	0	0	0	0	1	0	2	60	1	0	8	0	0	0	13	1	0	5	0	0	0	0	0	7	0	1	1844	1	7	1	0	8	0	0
55	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
56	0	4	0	0	0	0	0	0	0	2	0	0	3	0	4	0	0	0	0	2	0	0	0	0	5	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	2	1	0	14	0	0	8	0	0
57	0	3	0	0	0	11	0	0	0	1	0	0	0	0	4	3	0	0	0	7	7	0	0	0	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	31	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
59	0	27	0	0	0	0	0	0	0	11	0	0	9	0	18	1	0	0	0	4	2	0	0	0	24	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	18	0	5	0	0	147	0	0
60	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table B.2: Sokolov 2010, confusion matrix between flight-lines 2 and 3. Total matches: 124352, $\kappa = 0.67$, $\hat{\sigma}_\kappa = 0.0016$. See table 6.2 for classes index definitions.

Classes from Flight-line 4 \ Classes from Flight-line 3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	4872	0	0	0	19	0	10	0	83	0	30	3	1	24	0	0	0	16	71	33	0	6	0	69	0	0	0	0	4	2	0	0	5	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	4	0	0	9	0	0	2	0	3	0	0
3	0	0	120	0	21	0	0	0	0	0	0	0	0	0	32	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	2	5	0	36	0	0	0	0	0	0	0	1	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
6	0	35	0	0	0	168	1	0	0	0	0	0	0	0	6	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
7	0	0	0	0	0	0	42	19	4	0	0	0	4	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	26	0	0	0	0	9	633	28	176	0	6	24	1	22	0	0	0	0	1	0	0	0	0	3	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	31	0	0	0	0	0	0	0
9	0	6	0	0	1	0	1	67	151	1	0	0	13	0	73	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0
10	0	123	0	0	0	0	3	142	1	4185	0	131	17	10	138	1	0	0	0	2	7	0	0	0	82	0	0	0	0	0	0	0	0	24	4	0	0	0	0	0	0	8	0	0	5	0	0	0	0	0	6	0	0	140	1	0	1	0	4	3	0
11	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	29	0	0	0	0	0	1	0	82	0	325	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	26	0	0	0	0	3	86	15	36	0	2	126	0	68	0	0	0	0	1	1	0	0	0	6	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	1	0	0	0	0	0
14	0	14	0	0	0	0	0	0	0	2	0	3	4	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	86	3	0	20	1	0	111	72	203	0	0	100	2	4380	86	1	0	0	32	5	0	357	0	703	0	0	0	0	1	0	0	0	1	2	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	321	0	3	23	0	6	0	0
16	0	1	0	0	0	0	0	0	0	0	0	0	0	0	25	18	1	0	0	3	0	0	46	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	111	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1233	251	98	0	0	15	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	712	0	0	0	1	0	1	0	10	0	0	1	0	25	2	0	0	1192	11606	265	75	188	27	539	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	3	0	0	0	0	0
21	0	75	0	0	0	0	0	0	0	5	0	0	0	0	4	0	0	0	73	160	648	0	1	0	111	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	4	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	23	1	0	0	3	0	0	0	1	0	0	0	0	45	9	0	0	2	740	4	5	1122	0	255	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	7	0	0	0	0	0
24	0	80	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	30	22	3	0	0	16	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	209	0	0	0	0	1	2	0	136	0	0	5	0	550	9	0	0	11	933	216	0	628	2	3313	0	0	0	0	1	0	0	0	4	6	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	3	0	4	90	0	0	32	0	6	1	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	1	0	0	0	0	0	12	0	18	0	2	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	71	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0
35	0	2	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	1	0	7	0	0	0	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0
36	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
42	0	1	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
51	0	2	0	0	0	0	1	7	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0
52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
53	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0
54	0	50	0	0	0	0	0	144	16	258	0	0	10	0	212	0	0	0	0	9	3	0	2	0	124	0	0	0	0	0	0	0	0	3	25	0	0	0	0	0	0	14	0	0	8	0	0	0	0	0	0	0	1205	1	2	0	0	11	0	0
55	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	8	0	0	1	0	0	0
57	0	34	0	0	0	8	0	0	0	5	0	0	1	0	60	0	1	0	0	4	6	0	11	0	81	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	71	0	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
59	0	12	0	0	0	0	0	0	0	5	0	0	1	0	5	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	12	0	10	0	0	45	0	0
60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0
61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table B.3: Sokolov 2010, confusion matrix between flight-lines 3 and 4. Total matches: 49885, $\kappa = 0.66$, $\hat{\sigma}_\kappa = 0.0024$. See table 6.2 for classes index definitions.

| Classes from Flight-line 5 | Classes from Flight-line 4 | | |

Classes from Flight-line 6 \ Classes from Flight-line 5	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	2	0	0	0	0	0	0	0	76	0	0	0	0	4	0	0	0	0	2	3	0	0	0	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	2	0	0	0	0	0	0	0	3	0	0	0	0	764	18	0	0	0	3	1	0	33	0	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	189	47	1	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	61	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	218	1030	5	0	6	11	24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	37	8	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	30	0	0	26	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	13	0	0	0	0	0	0	0	6	0	0	0	0	8	0	0	0	7	131	11	0	68	0	105	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0
55	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

| Classes from Flight-line 7 | Classes from Flight-line 6 | | |

North-Eastern Classification \ South-Eastern Classification																															
Classes	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	$1.6E^5$	16	32	0	216	0	224	32	$1.2E^4$	0	2388	1068	88	2048	4	0	0	3408	$1.0E^4$	7254	0	464	648	9112	0	0	0	64	96	208
3	0	96	$2.4E^4$	2608	888	0	16	512	0	256	0	32	0	0	9296	2896	0	0	0	0	0	0	288	0	136	0	0	0	0	0	0
4	0	176	1040	5424	0	0	0	240	0	160	0	0	0	0	2720	64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	16	1016	48	$1.1E^4$	0	0	0	448	0	0	0	8	0	5568	512	0	0	0	0	0	0	128	0	0	0	0	0	0	0	0
6	0	312	0	0	0	2224	8	0	0	0	0	0	68	0	80	0	0	0	0	8	0	0	8	0	0	0	0	0	0	0	0
7	0	4	0	0	0	0	6504	1864	32	80	0	32	2120	0	568	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	880	32	32	96	0	2488	$9.0E^4$	2788	$2.4E^4$	16	2176	$1.1E^4$	200	6808	0	0	0	64	12	0	0	4	0	188	0	0	0	0	0	0
9	0	80	0	0	6376	0	76	1160	$1.5E^4$	8	0	0	432	0	6328	0	0	0	0	0	0	0	0	0	140	0	0	0	0	0	0
10	0	$2.3E^4$	32	0	0	0	88	$1.9E^4$	12	$1.1E^6$	256	4588	4092	832	$3.3E^4$	8	0	0	100	1700	2306	0	112	0	$3.6E^4$	0	0	0	0	0	0
11	0	208	0	0	0	0	8	8	8	624	112	64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	2696	0	0	0	0	0	2136	0	7720	64	$2.2E^4$	1472	56	72	0	0	0	0	32	16	0	0	0	48	0	0	0	0	0	36
13	0	2204	0	0	32	128	248	3624	348	2948	0	688	$4.2E^4$	0	9720	0	0	0	0	100	344	0	36	0	1456	0	0	0	0	0	0
14	0	176	0	0	0	0	0	0	0	272	0	56	32	184	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	5232	2968	1824	5856	104	320	3760	1576	$1.1E^4$	0	96	4640	96	$3.3E^5$	$2.5E^4$	184	0	48	5118	1034	48	$3.7E^4$	0	$9.0E^4$	0	0	0	0	8	0
16	0	124	832	32	112	0	0	0	0	0	0	0	32	0	2328	6686	56	0	0	744	16	0	6244	0	990	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	80	0	0	0	0	0	96	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	$3.2E^4$	0	0	0	0	0	0	0	40	0	0	0	0	48	0	0	0	$4.0E^5$	$6.0E^4$	6658	64	64	3626	1524	0	0	0	0	0	0
20	0	$7.5E^4$	0	0	0	8	0	8	0	944	0	32	100	0	1514	208	0	0	$2.2E^5$	$1.2E^6$	$2.0E^4$	3833	$1.5E^4$	4611	$3.8E^4$	0	0	0	0	0	0
21	0	$2.2E^4$	0	0	0	0	0	0	0	2424	0	96	96	0	660	16	0	0	$8.4E^4$	$7.8E^4$	$1.2E^5$	0	248	1136	$2.10E^4$	0	0	0	0	0	0
22	0	32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	288	0	160	32	0	0	0	0	0	0	0	0
23	0	1482	136	0	32	24	0	0	0	72	0	0	36	0	1760	1392	80	0	228	$6.4E^4$	160	732	$5.8E^4$	64	7164	0	0	0	0	0	0
24	0	$1.2E^4$	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	2655	1292	56	0	0	2677	48	0	0	0	0	0	0
25	0	$3.1E^4$	48	0	0	0	40	16	0	$1.7E^4$	0	32	812	0	$3.2E^4$	1692	0	0	5702	$1.3E^5$	$3.9E^4$	32	$5.6E^4$	224	$3.6E^5$	0	0	0	0	8	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	64	0	0	0	0	0	0	0	0	0	32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32	0	0
30	0	120	0	0	0	0	0	0	0	32	0	0	0	0	0	0	0	0	0	0	96	0	0	0	40	0	0	0	0	0	0
31	0	576	0	0	0	0	0	0	0	0	0	128	0	0	0	0	0	0	0	0	64	0	0	0	32	0	0	0	0	32	288
32	0	160	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	2392	0	32	0	0	0	160	0	4512	0	80	64	56	128	0	0	0	0	0	0	0	0	0	80	0	0	0	0	0	0
35	0	192	288	0	80	0	0	0	0	32	0	0	32	0	1200	496	0	0	0	80	64	0	232	0	1640	0	0	0	0	0	0
36	0	288	0	0	0	16	0	0	0	64	0	0	64	0	160	0	0	0	0	64	64	0	0	0	192	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	8	0	0	0	0	0	0	0	64	0	0	0	0	0	0	0	0	0	32	64	0	0	0	144	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	0	544	0	0	0	0	0	0	0	32	0	0	0	0	0	0	0	0	0	32	64	0	0	0	0	0	0	0	0	0	0
42	0	776	0	0	0	0	0	0	0	1136	0	0	0	0	0	0	0	0	0	0	0	0	16	0	64	0	0	0	0	0	0
43	0	0	16	0	0	0	0	32	0	128	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32	0	0	0	0	0	0
44	0	96	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32	0	0	0	0	0	0	0	0	0	0
45	0	240	0	0	0	0	0	0	0	128	0	0	0	0	32	0	0	0	0	0	0	0	0	0	144	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	0	64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32	0	0	0	0	0	0	0	0
48	0	64	0	0	0	0	0	0	0	64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
51	0	512	0	0	0	0	8	56	0	1040	0	0	128	0	128	0	0	0	0	16	32	0	0	0	408	0	0	0	0	0	0
52	0	96	0	0	0	0	0	0	0	32	0	0	16	0	32	0	0	0	0	0	64	0	0	0	0	0	0	0	0	0	0
53	0	24	0	0	0	0	0	0	0	0	0	0	0	0	64	0	0	0	0	64	8	0	0	0	120	0	0	0	0	0	0
54	0	5152	384	384	0	0	32	3776	128	$1.5E^4$	0	80	368	0	$2.1E^4$	0	0	0	48	858	552	0	160	0	$2.0E^4$	0	0	0	0	0	0
55	0	16	0	0	0	0	0	128	0	320	0	0	0	0	0	0	0	0	0	8	8	0	32	0	64	0	0	0	0	0	0
56	0	64	0	0	0	0	0	0	0	32	0	0	48	0	96	0	0	0	0	32	0	0	0	0	272	0	0	0	0	0	0
57	0	1256	0	0	0	240	0	0	0	288	0	32	328	0	2348	144	8	0	0	1092	896	0	668	0	5360	0	0	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
59	0	1360	0	0	0	0	0	0	0	888	0	0	216	0	968	16	0	0	64	160	64	0	0	32	1104	0	0	0	0	0	0
60	0	48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32	0	0	0	0	0	0
61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table B.7: Sokolov 2010, confusion matrix between south-east and north-west flight lines - first part. Total matches: 608756, $\kappa = 0.63$ $\hat{\sigma}_\kappa = 5.8\ 10^{-4}$. See table 6.2 for classes index definitions.

North-Eastern Classification \ South-Eastern Classification	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	112	0	2568	48	8	0	128	0	0	192	1104	0	32	444	0	0	0	0	0	352	0	16	2680	0	16	1156	0	472	0	0
3	0	0	0	1472	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1408	0	0	0	0	0	0	0
4	0	0	0	96	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1008	0	0	0	0	0	0	0
5	0	0	0	640	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	88	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	48	0	0	0	0	32	0	0
8	32	0	160	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	152	0	0	6536	0	0	0	0	32	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	256	0	0	0	0	0	0	0
10	32	0	3168	400	160	0	256	0	0	32	3296	96	0	392	0	0	0	0	0	960	0	0	$4.4E^4$	24	320	362	0	1744	120	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	16	0	32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32	0	0	64	0	0	0	0
13	0	0	8	0	32	0	0	0	0	0	16	0	0	0	0	0	0	0	0	144	0	0	104	0	0	1508	0	80	0	0
14	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0
15	0	0	104	3360	120	0	32	0	0	0	48	32	0	48	0	0	0	0	0	16	32	224	$1.6E^4$	16	232	6472	0	720	0	0
16	0	0	0	84	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	72	32	0	0	8	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0
20	0	0	160	128	16	0	80	0	0	0	0	0	0	0	0	0	0	0	0	16	0	16	492	0	32	552	0	16	0	0
21	0	0	64	96	64	0	48	0	0	0	32	32	64	0	0	0	0	0	0	32	0	0	868	0	32	434	0	128	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	64	0	0	0	0	0	0	0	0	0	0	24	0	0	504	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	192	2352	224	0	312	0	0	0	400	32	0	80	0	0	0	32	0	168	0	80	$1.5E^4$	0	400	4596	0	1280	8	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	0	$1.2E^4$	0	0	0	0	0	0	32	352	0	0	32	0	0	0	32	0	280	0	0	1376	0	0	0	0	0	0	0
35	0	0	0	3868	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	312	0	32	0	0	32	0	0
36	0	0	32	0	224	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	64	0	0	0	0	72	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	152	0	0	0	0	0	0	0
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	0	0	64	0	0	0	0	0	0	96	0	0	0	0	0	0	0	0	0	64	0	0	0	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0	0	0	0	1392	0	0	96	0	0	0	0	0	0	0	0	1112	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	0	0	0	64	0	0	0	0	0	0	0	0	0	0	160	0	0	0	0	0	0	0
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	0	0	0	96	0	0	0	0	0	0	96	0	0	800	0	0	0	0	0	0	0	0	752	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	384	0	0	0	0	0	0	64	448	0	0	0	0	0	0
48	0	0	32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
51	0	0	424	32	32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1400	16	0	192	0	32	0	0	288	0	0
52	0	0	64	0	0	0	0	0	0	0	32	0	0	0	0	0	0	0	0	32	16	0	0	0	0	0	0	0	0	0
53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	80	0	0	32	0	0	0	0	0
54	16	0	216	1544	16	0	192	0	0	0	736	48	0	272	0	0	0	32	0	112	0	16	$1.0E^5$	24	128	80	0	216	288	0
55	0	0	320	0	0	0	0	0	0	0	0	160	0	0	0	64	0	0	0	0	32	0	384	1024	0	0	0	0	0	0
56	0	0	0	200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	64	0	64	40	0	960	0	0	360	0	0
57	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	48	0	0	8372	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
59	0	0	0	112	248	0	0	0	0	0	0	0	0	0	0	0	0	0	0	88	0	128	576	0	512	0	0	5560	0	0
60	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	64	0	0	0	0	0	64	0
61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table B.8: Sokolov 2010, confusion matrix between south-east and north-west flight lines - second part. Total matches: 608756, $\kappa = 0.63$ $\hat{\sigma}_\kappa = 5.8\ 10^{-4}$. See table 6.2 for classes index definitions.

B.1.2 Proportions of Asymmetries and Symmetries

Flight-Lines 1 & 2

Class	p_u	p_p	p_a	p_s	Class	p_u	p_p	p_a	p_s
1	0.638	0.334	0.304	0.972	2	0.530	0.762	−0.232	1.292
3	0.579	0.553	0.025	1.132	4	0.705	0.466	0.238	1.171
5	0.556	0.385	0.171	0.940	6	0.589	0.664	−0.075	1.253
7	0.617	0.748	−0.131	1.365	8	0.512	0.765	−0.253	1.277
9	0.745	0.829	−0.084	1.574	10	N/A	0.000	N/A	N/A
11	0.734	0.753	−0.018	1.487	12	0.673	0.631	0.042	1.303
13	N/A	0.000	N/A	N/A	14	0.603	0.717	−0.113	1.320
15	0.274	0.139	0.135	0.413	16	0.500	0.167	0.333	0.667
18	0.793	0.537	0.256	1.330	19	0.754	0.774	−0.021	1.528
20	0.387	0.626	−0.239	1.013	21	0.385	0.079	0.305	0.464
22	0.443	0.292	0.151	0.736	23	0.140	0.269	−0.130	0.409
24	0.529	0.608	−0.078	1.137	28	0.250	0.333	−0.083	0.583
29	0.000	0.000	0.000	0.000	30	0.257	0.562	−0.305	0.820
31	0.000	N/A	N/A	N/A	33	0.578	0.634	−0.057	1.212
34	0.466	0.343	0.123	0.809	35	0.182	0.200	−0.018	0.382
37	0.000	0.000	0.000	0.000	40	0.107	0.273	−0.166	0.380
41	0.348	0.191	0.156	0.539	42	0.167	0.133	0.033	0.300
43	0.000	0.000	0.000	0.000	44	0.265	0.231	0.034	0.495
46	0.387	0.857	−0.470	1.244	47	0.000	N/A	N/A	N/A
48	N/A	0.000	N/A	N/A	50	0.225	0.333	−0.108	0.558
51	0.000	0.000	0.000	0.000	52	0.000	0.000	0.000	0.000
53	0.636	0.498	0.138	1.134	54	0.405	0.696	−0.291	1.101
55	0.583	0.382	0.202	0.965	56	0.422	0.348	0.073	0.770
58	0.408	0.494	−0.086	0.903	59	0.500	0.143	0.357	0.643

Table B.9: Symmetric and asymmetric accuracies: p_s and p_a along with so-called producer's and user's accuracies (respectively p_p and p_u) for classifications results between flight-lines 1 and 2: Sokolov 2010 dataset. Classes with absolutely no matches were removed from the table.

Flight-Lines 2 & 3

Class	p_u	p_p	p_a	p_s	Class	p_u	p_p	p_a	p_s
1	0.754	0.416	0.338	1.170	2	0.548	0.780	−0.232	1.328
3	0.384	0.464	−0.080	0.849	4	0.081	0.215	−0.134	0.296
5	0.900	0.776	0.124	1.676	6	0.062	0.200	−0.138	0.263
7	0.559	0.598	−0.040	1.157	8	0.312	0.714	−0.402	1.027
9	0.925	0.955	−0.029	1.880	10	0.111	0.292	−0.181	0.403
11	0.356	0.514	−0.158	0.870	12	0.189	0.211	−0.021	0.400
13	0.207	0.098	0.109	0.305	14	0.488	0.603	−0.115	1.090
15	0.517	0.213	0.304	0.729	16	0.429	0.250	0.179	0.679
18	0.802	0.624	0.177	1.426	19	0.771	0.750	0.021	1.521
20	0.179	0.491	−0.312	0.670	21	0.000	0.000	0.000	0.000
22	0.363	0.346	0.017	0.710	23	0.182	0.120	0.062	0.302
24	0.445	0.538	−0.093	0.984	31	0.300	0.188	0.112	0.487
33	0.448	0.541	−0.093	0.989	34	0.427	0.197	0.230	0.625
35	0.143	0.286	−0.143	0.429	37	0.000	0.000	0.000	0.000
41	0.091	0.108	−0.017	0.199	42	0.000	0.000	0.000	0.000
44	0.425	0.633	−0.208	1.057	50	0.388	0.398	−0.010	0.787
51	0.167	0.500	−0.333	0.667	52	0.500	0.056	0.444	0.556
53	0.510	0.563	−0.053	1.074	54	0.000	0.000	0.000	0.000
55	0.230	0.286	−0.056	0.515	56	0.320	0.242	0.077	0.562
58	0.542	0.505	0.037	1.048	59	0.000	N/A	N/A	N/A

Table B.10: Symmetric and asymmetric accuracies: p_s and p_a along with the so-called producer's and user's accuracies (respectively p_p and p_u) for classification results between flight-lines 2 and 3: Sokolov 2010 data-set. Classes with absolutely no matches were removed from the table.

Flight-Lines 3 & 4

Class	p_u	p_p	p_a	p_s	Class	p_u	p_p	p_a	p_s
1	0.925	0.745	0.180	1.670	2	0.690	0.930	−0.241	1.620
4	0.424	0.462	−0.038	0.885	5	0.789	0.832	−0.043	1.620
6	0.583	0.677	−0.094	1.261	7	0.649	0.513	0.137	1.162
8	0.448	0.524	−0.076	0.972	9	0.831	0.803	0.028	1.633
10	0.000	N/A	N/A	N/A	11	0.734	0.651	0.082	1.385
12	0.334	0.406	−0.072	0.741	13	0.306	0.407	−0.102	0.713
14	0.671	0.765	−0.094	1.437	15	0.154	0.144	0.010	0.298
16	0.000	0.000	0.000	0.000	18	0.719	0.482	0.237	1.201
19	0.792	0.839	−0.047	1.631	20	0.597	0.502	0.095	1.099
21	0.000	0.000	0.000	0.000	22	0.506	0.475	0.031	0.981
23	0.104	0.267	−0.163	0.371	24	0.537	0.621	−0.084	1.158
29	0.000	0.000	0.000	0.000	30	N/A	0.000	N/A	N/A
33	0.634	0.538	0.096	1.172	34	0.475	0.322	0.153	0.797
35	0.000	0.000	0.000	0.000	37	0.000	0.000	0.000	0.000
41	0.231	0.214	0.016	0.445	44	0.500	0.125	0.375	0.625
50	0.290	0.243	0.047	0.534	52	0.615	0.533	0.082	1.149
53	0.575	0.642	−0.068	1.217	54	0.000	0.000	0.000	0.000
55	0.615	0.348	0.268	0.963	56	0.250	0.486	−0.236	0.736
58	0.459	0.584	−0.125	1.044	59	0.000	0.000	0.000	0.000

Table B.11: Symmetric and asymmetric accuracies: p_s and p_a along with the so-called producer's and user's accuracies (respectively p_p and p_u) for classification results between flight-lines 3 and 4: Sokolov 2010 data-set. Classes with absolutely no matches were removed from the table.

Flight-lines 4 & 5

Class	p_u	p_p	p_a	p_s	Class	p_u	p_p	p_a	p_s
1	0.509	0.166	0.344	0.675	5	0.000	N/A	N/A	N/A
6	0.400	0.667	−0.267	1.067	7	0.534	0.633	−0.098	1.167
8	0.486	0.773	−0.287	1.258	9	0.879	0.906	−0.027	1.785
11	0.222	0.250	−0.028	0.472	12	0.405	0.451	−0.046	0.856
14	0.416	0.761	−0.345	1.177	15	0.143	0.167	−0.024	0.310
18	0.798	0.377	0.421	1.175	19	0.789	0.815	−0.026	1.604
20	0.387	0.580	−0.193	0.967	21	N/A	0.000	N/A	N/A
22	0.484	0.492	−0.008	0.976	23	0.085	0.135	−0.050	0.220
24	0.512	0.658	−0.146	1.170	30	N/A	0.000	N/A	N/A
34	1.000	0.500	0.500	1.500	37	0.000	N/A	N/A	N/A
44	N/A	0.000	N/A	N/A	53	0.286	0.154	0.132	0.440
56	0.500	0.352	0.148	0.852					

Table B.12: Symmetric and asymmetric accuracies: p_s and p_a along with the so-called producer's and user's accuracies (respectively p_p and p_u) for classification results between flight-lines 4 and 5: Sokolov 2010 data-set. Classes with absolutely no matches were removed from the table.

Flight-lines 5 & 6

Class	p_u	p_p	p_a	p_s	Class	p_u	p_p	p_a	p_s
1	0.375	0.027	0.348	0.402	9	0.685	0.894	−0.209	1.579
12	0.000	N/A	N/A	N/A	14	0.898	0.978	−0.080	1.876
15	0.750	0.143	0.607	0.893	18	0.721	0.442	0.280	1.163
19	0.760	0.803	−0.043	1.562	20	0.101	0.250	−0.149	0.351
22	0.426	0.195	0.231	0.622	23	0.000	0.000	0.000	0.000
24	0.299	0.522	−0.223	0.822	53	0.529	0.900	−0.371	1.429
56	N/A	0.000	N/A	N/A					

Table B.13: Symmetric and asymmetric accuracies: p_s and p_a along with the so-called producer's and user's accuracies (respectively p_p and p_u) for classification results between flight-lines 5 and 6: Sokolov 2010 data-set. Classes with absolutely no matches were removed from the table.

Fligth-lines 6 & 7

Class	p_u	p_p	p_a	p_s	**Class**	p_u	p_p	p_a	p_s
1	1.000	0.759	0.241	1.759	14	0.000	N/A	N/A	N/A
18	0.780	0.597	0.182	1.377	19	0.755	0.925	-0.170	1.680
21	N/A	0.000	N/A	N/A	22	N/A	0.000	N/A	N/A
23	0.500	0.167	0.333	0.667	24	0.476	0.714	-0.238	1.190
53	0.250	0.250	0.000	0.500					

Table B.14: Symmetric and asymmetric accuracies: p_s and p_a along with the so-called producer's and user's accuracies (respectively p_p and p_u) for classification results between flight-lines 6 and 7: Sokolov 2010 data-set. Classes with absolutely no matches were removed from the table.

B.2 Parameter Optimisation

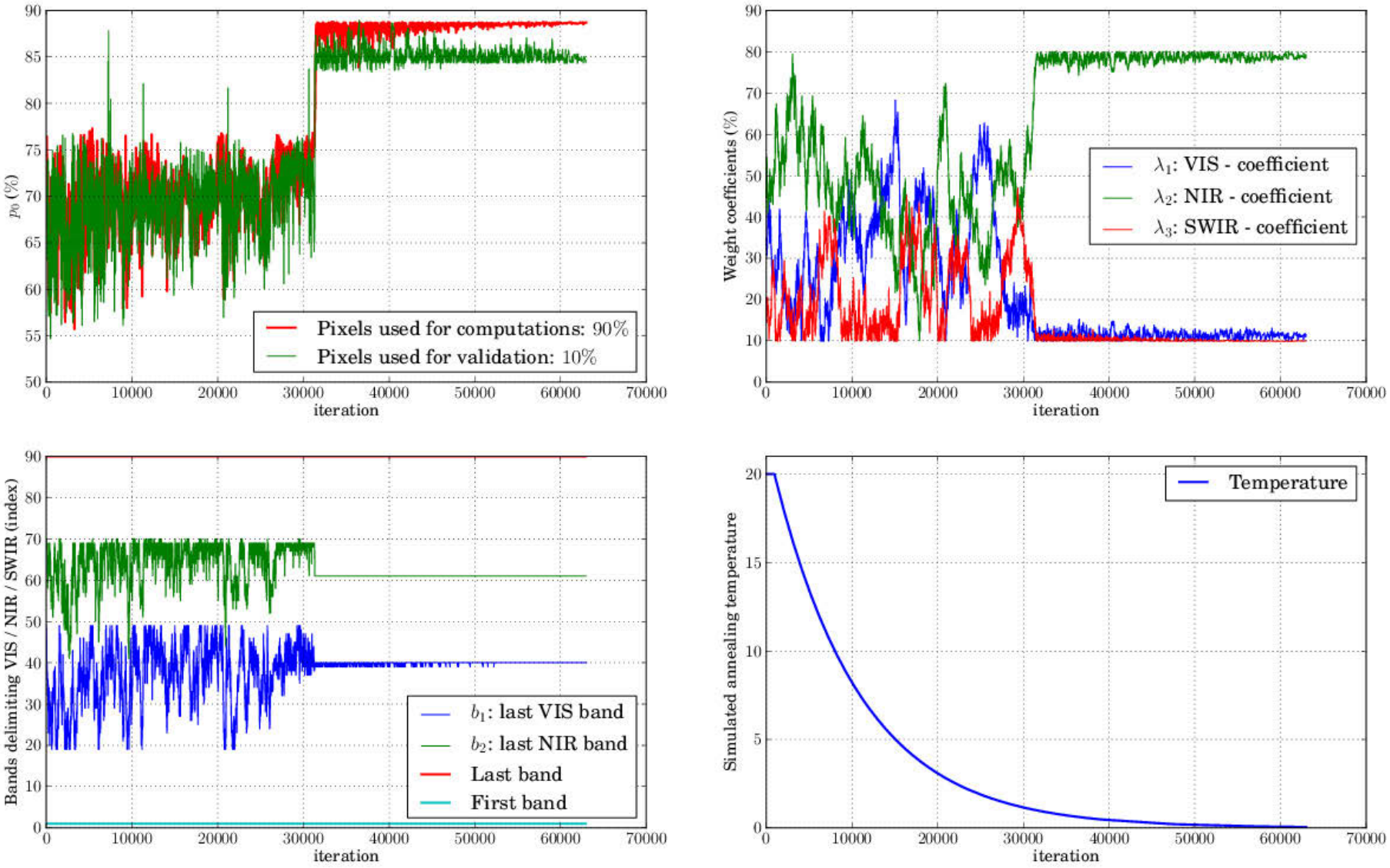

Figure B.1: Simulated annealing, evolution of the first run.

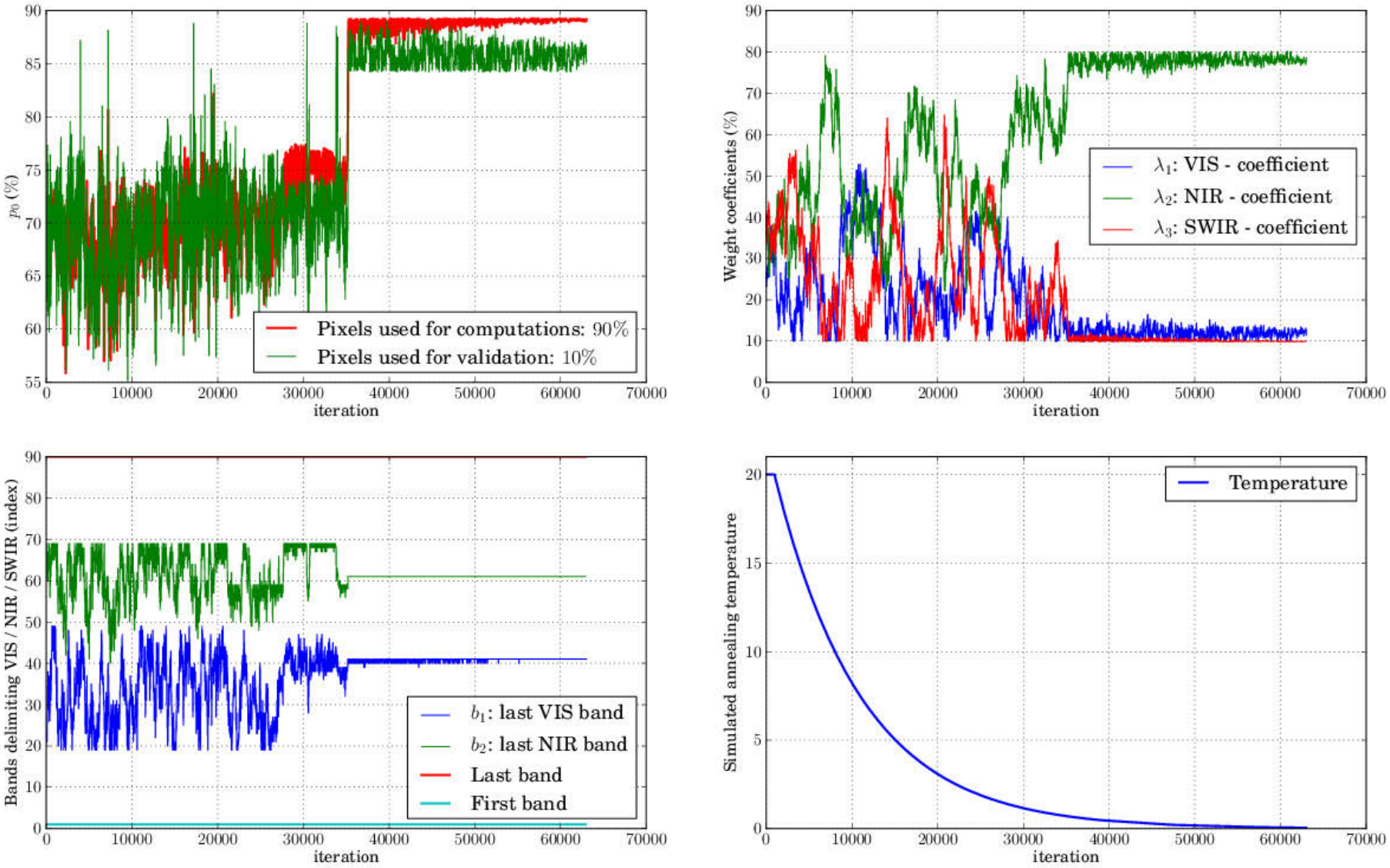

Figure B.2: Simulated annealing, evolution of the second run.

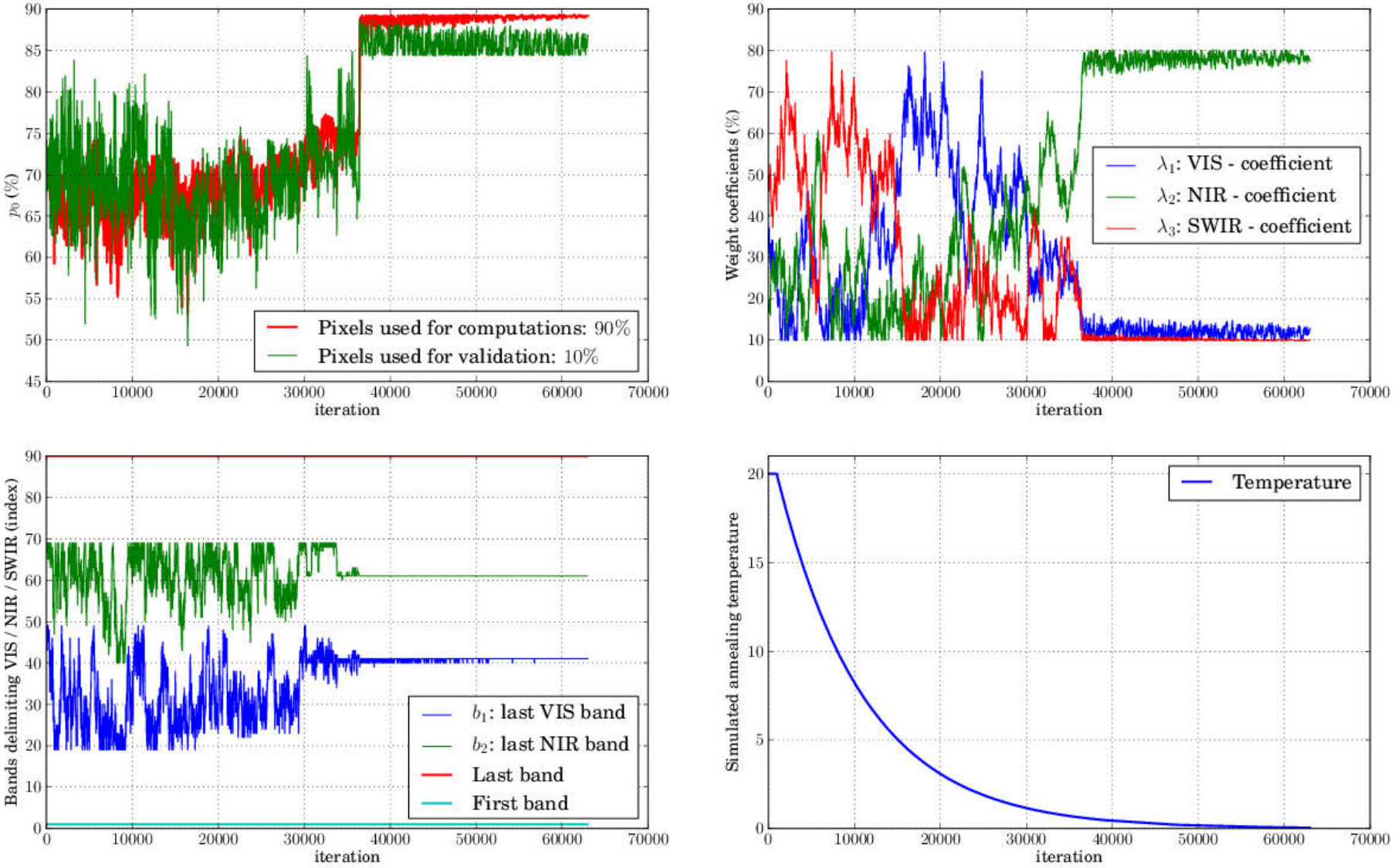

Figure B.3: Simulated annealing, evolution of the third run.

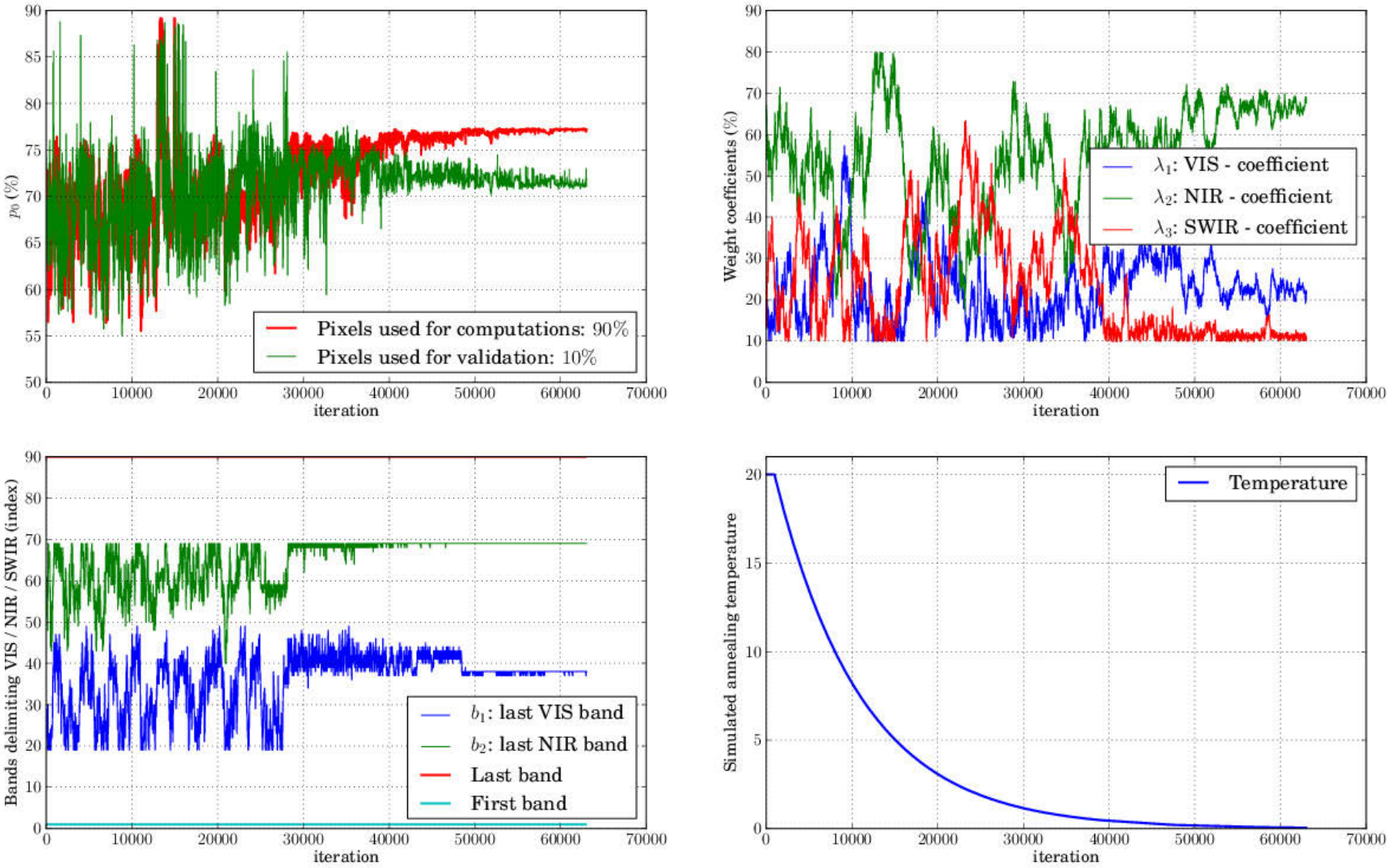

Figure B.4: Simulated annealing, evolution of the fourth run.

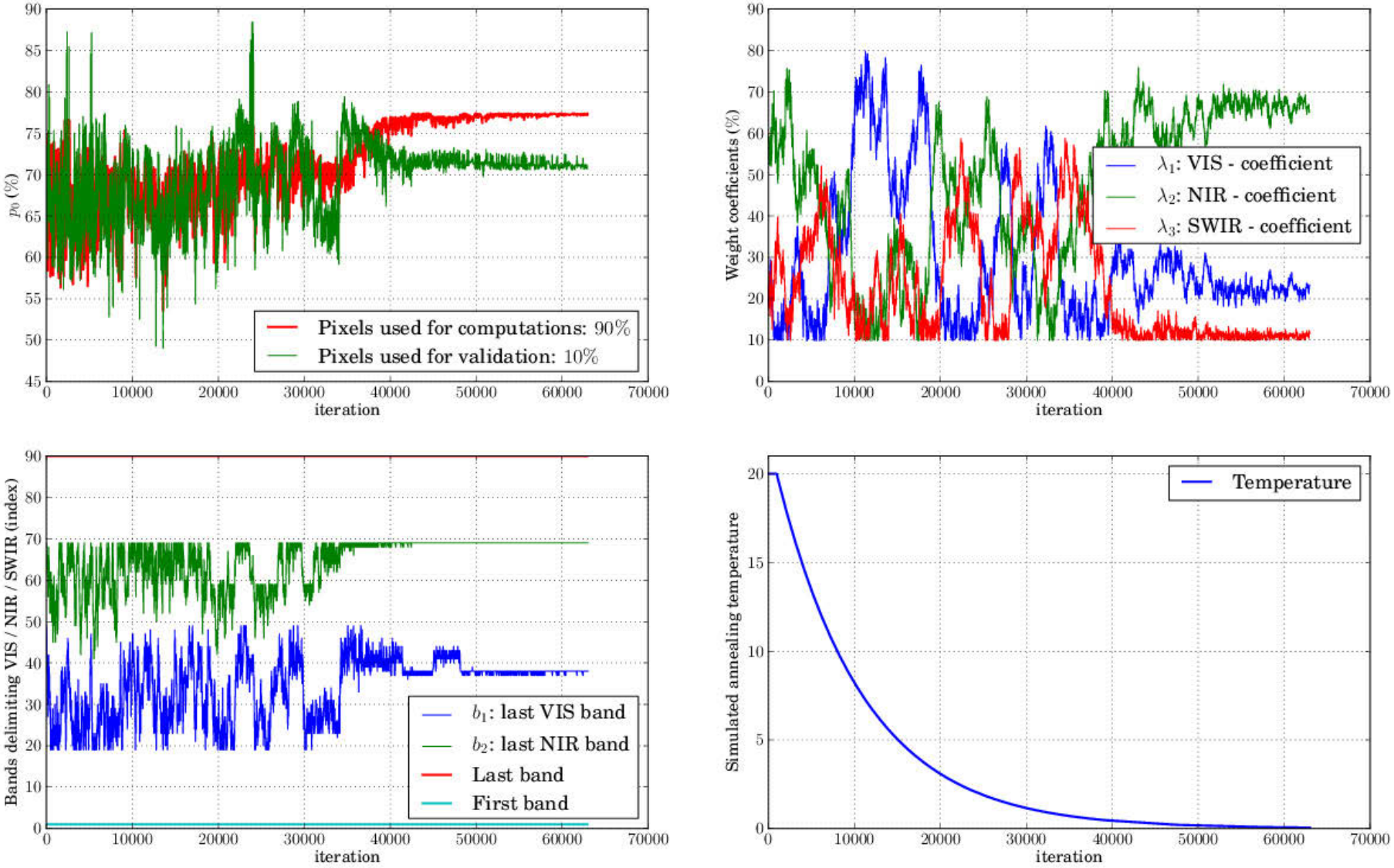

Figure B.5: Simulated annealing, evolution of the fifth run.

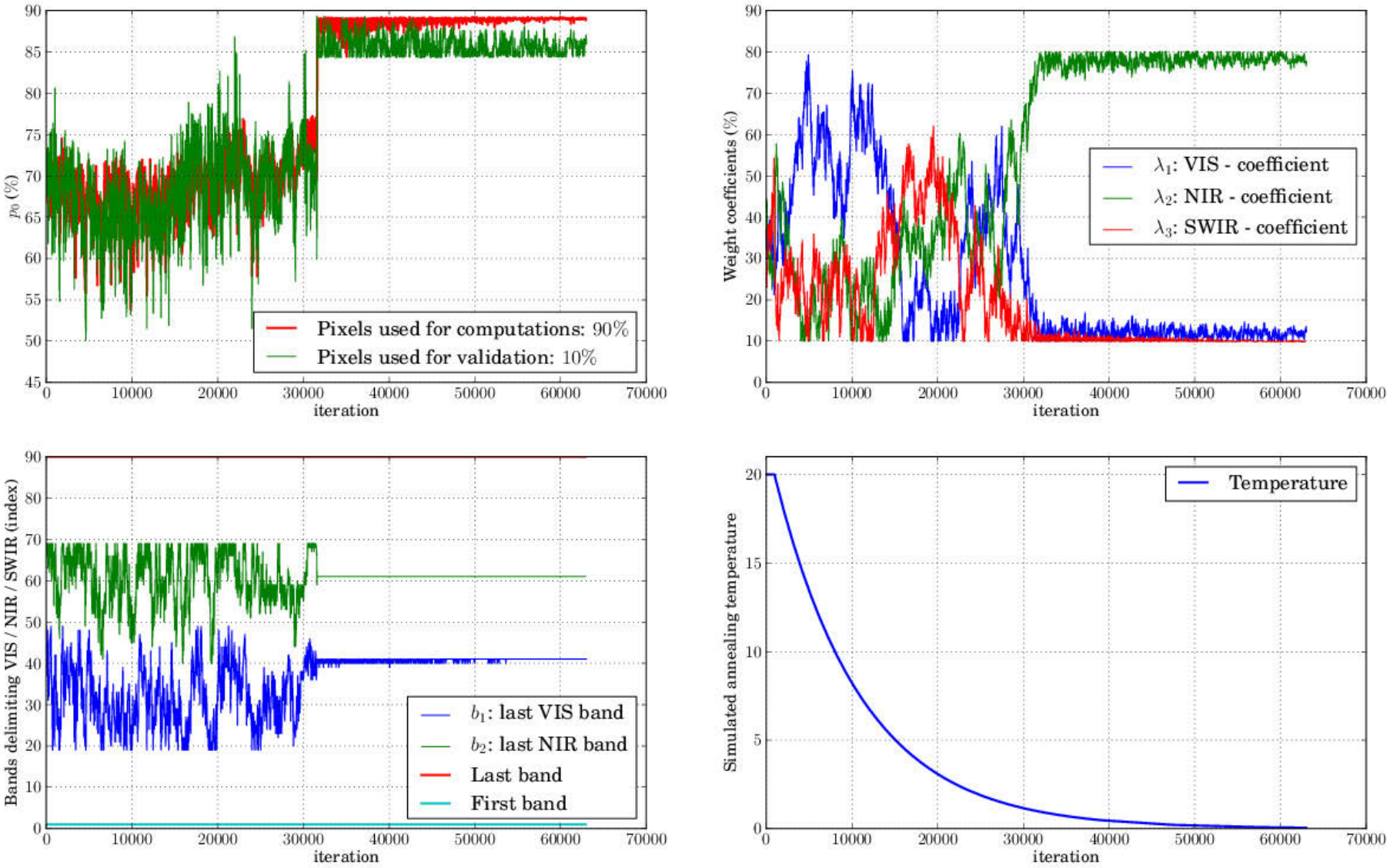

Figure B.6: Simulated annealing, evolution of the sixth run.

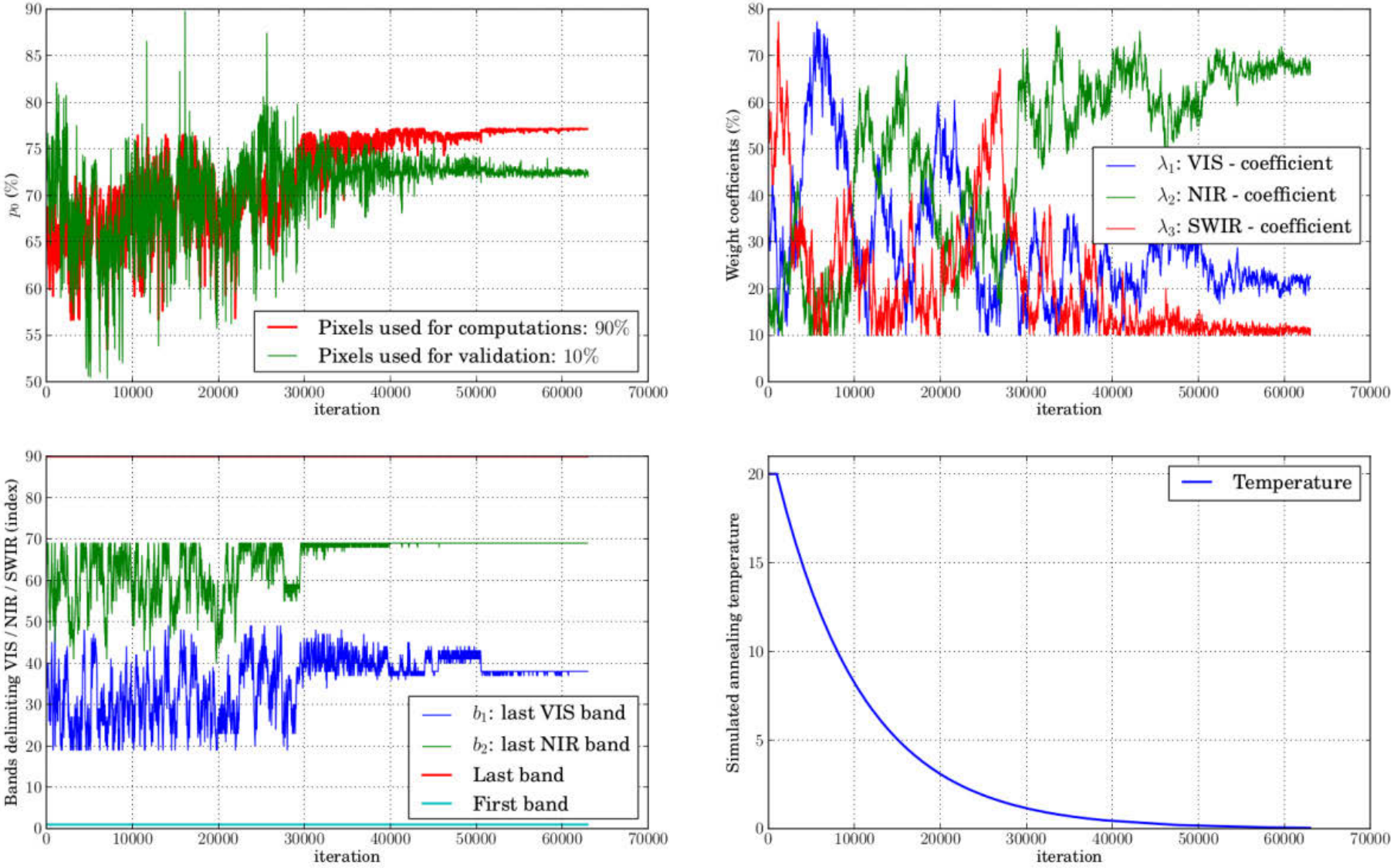

Figure B.7: Simulated annealing, evolution of the seventh run.

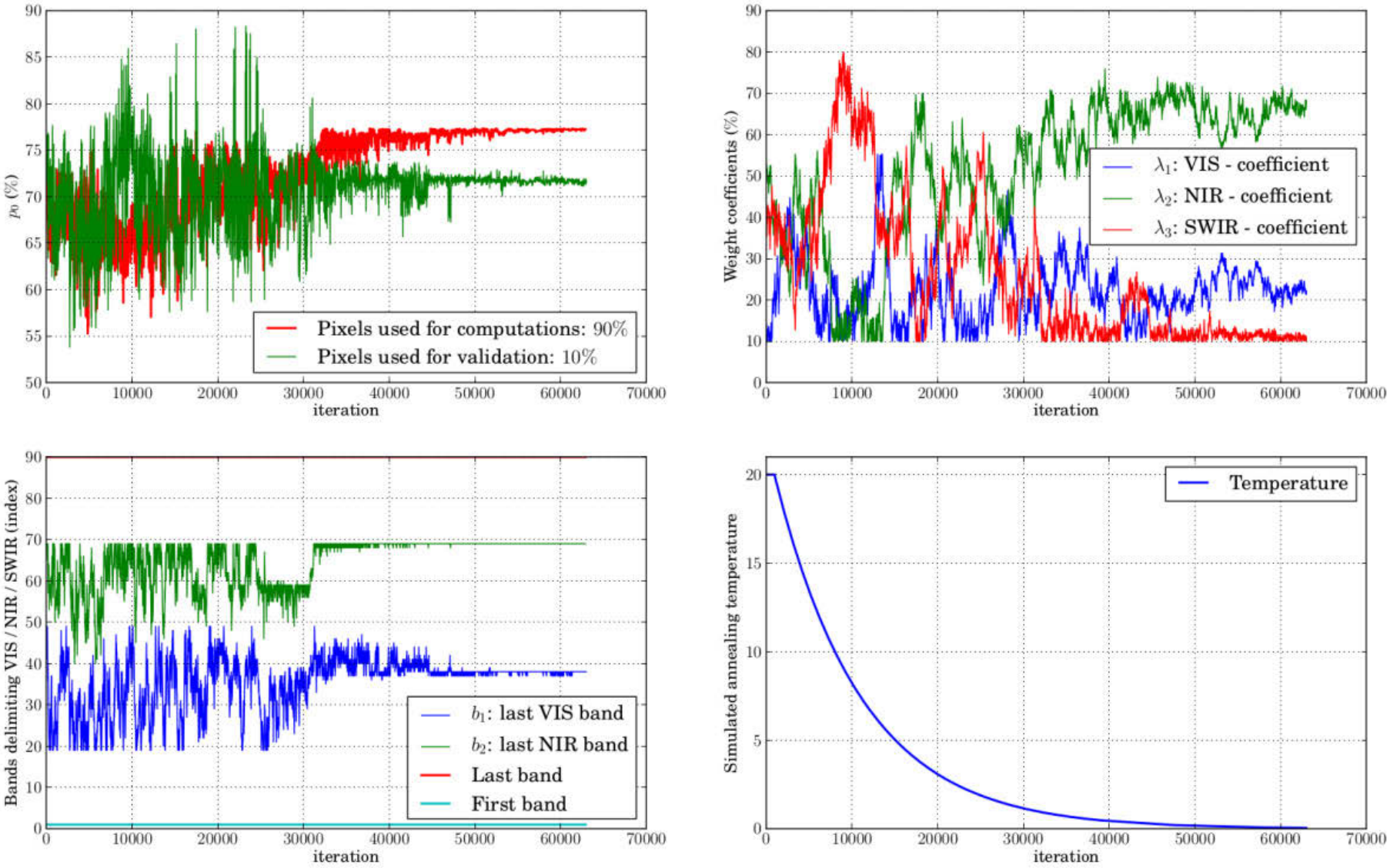

Figure B.8: Simulated annealing, evolution of the eigthth run.

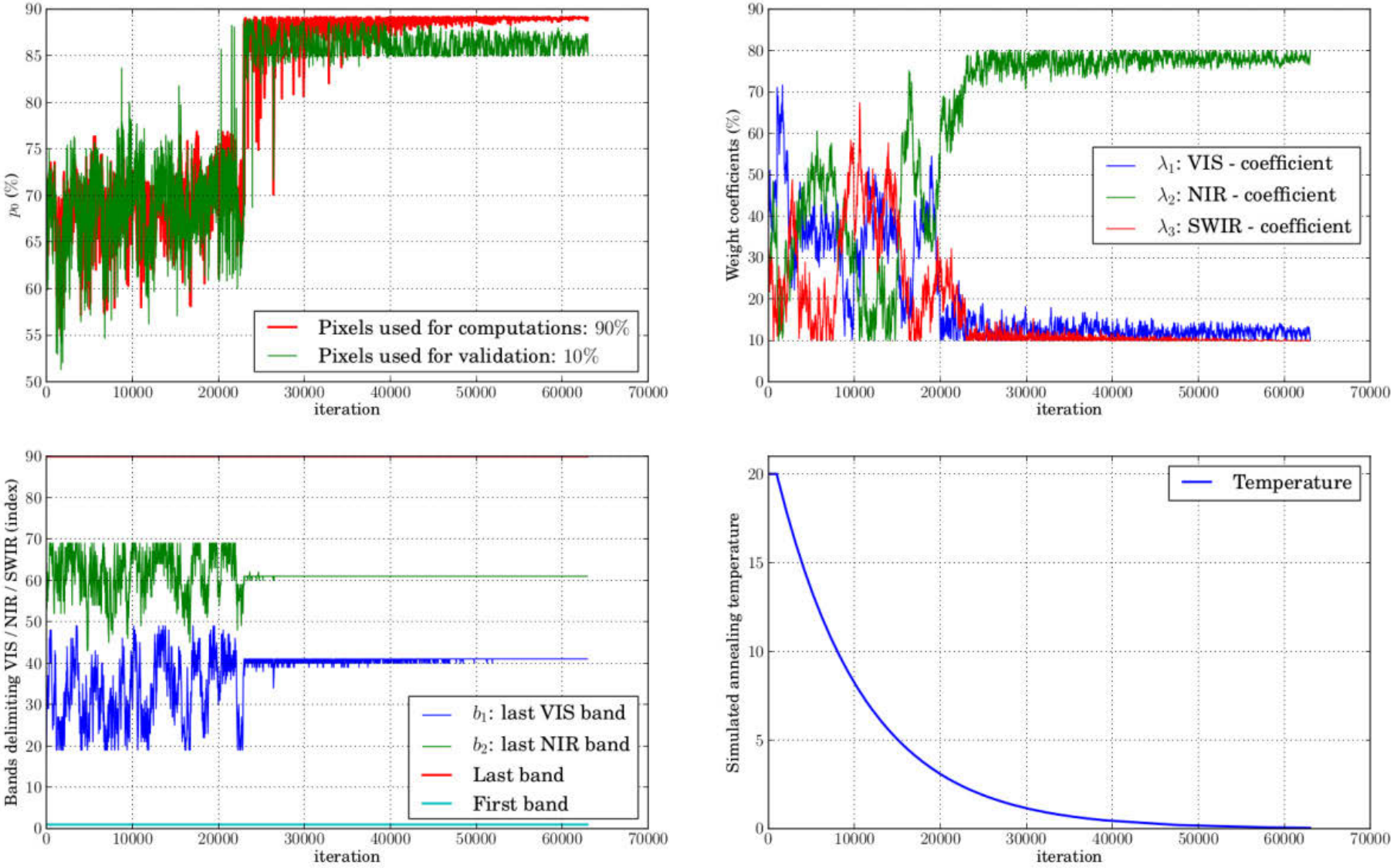

Figure B.9: Simulated annealing, evolution of the nineth run.

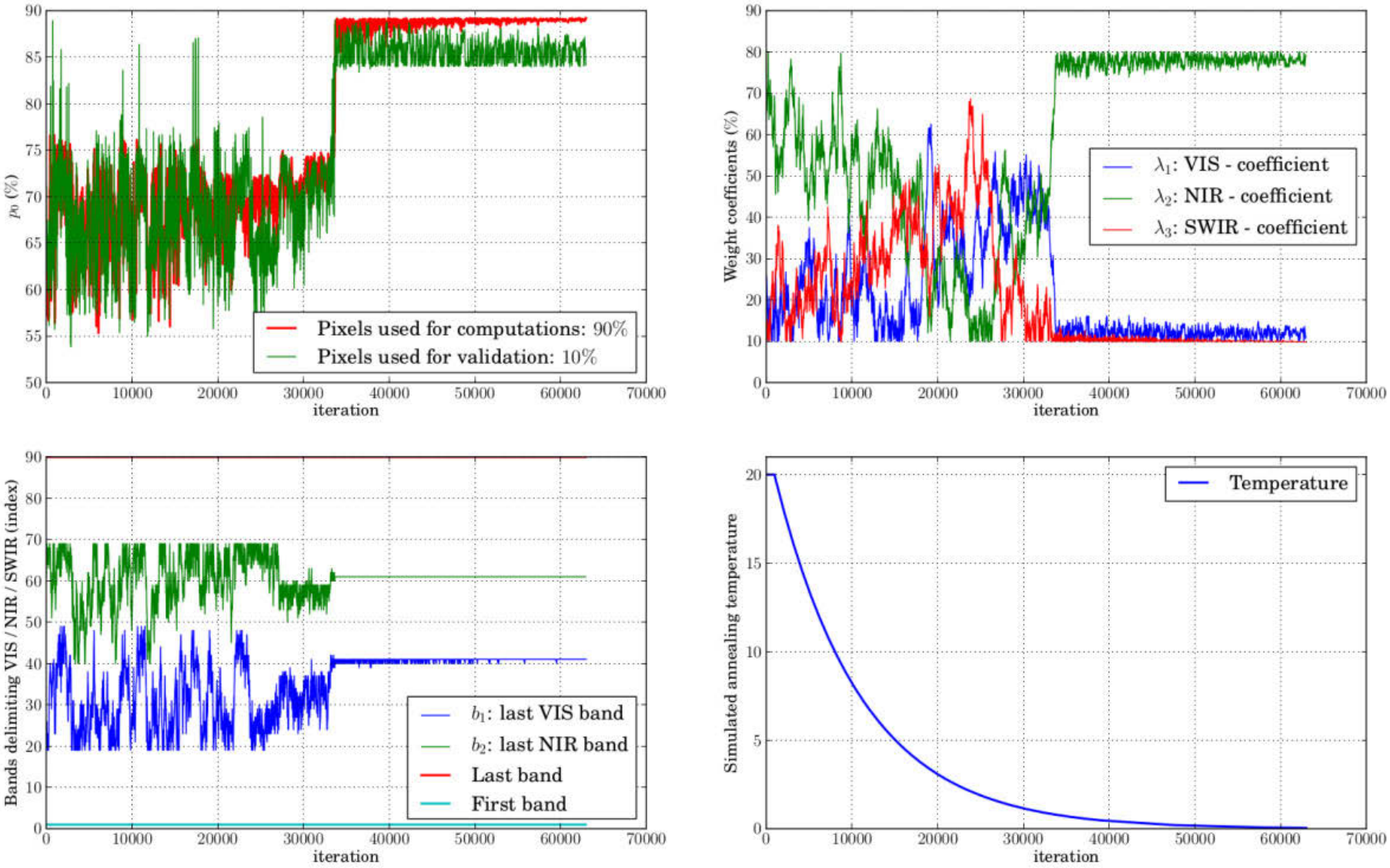

Figure B.10: Simulated annealing, evolution of the tenth run.

Appendix C

Elements of Statistics

> The most important questions of life are indeed, for the most part, really only problems of probability.
>
> Pierre-Simon Laplace - Scientist

C.1 Basic Statistics

C.1.1 Definitions

This section reviews some basic definitions related to probabilities and density of probabilities. Again, main ideas and concepts are presented but without focusing on the underlying mathematical details.

The probability density function, or PDF, can be seen as the limit case of a histogram. It is a positive function such as the probability that a random variable is taking values between a and b is equal to the integral of the PDF between both values. As a result, a PDF has a normalised integral equal to one.

The probability density function is strongly bound with the characteristic function obtained by an inverse Fourier transform of the PDF. A similar function is the cumulative density function corresponding to the integral. In practice the value of the cumulative distribution at x represents the probability that a random variable is smaller than x.

C.1.2 Maximum Likelihood Estimation

The maximum likelihood estimation aims at guessing the parameters defining a statistical model, given a set of N samples: $\{x_k\}_k$. The underlying idea consists of making use of the Bayes relation for obtaining the parameters that that are the most likely given the N observed samples.

Writing θ the model's parameter, the goal is to maximise: $P(\theta|\{x_k\}_k)$. Inverting this relationship by means of the Bayes relation, this yields equation C.1.1.

$$LE(\theta) = P(\theta|\{x_k\}_k) = P(\{x_k\}_k|\theta)P(\theta)/P(\{x_k\}_k) \propto P(\{x_k\}_k|\theta)P(\theta) \tag{C.1.1}$$

Assuming that the parameters are equiprobable finally removes the dependency on $P(\theta)$, thus allowing to perform the maximisation on $P(\{x_k\}_k|\theta)$ directly. This can furthermore be rewritten as

$$\arg\max_{\theta} LE\left(\theta\right) = \arg\max_{\theta} \prod_{k=1}^{N} P(x_k|\theta)$$

Since the statistical model is known, $P(x_k|\theta)$ can be estimated for any given value of θ, thus leading the path to an estimation of the parameters.

C.2 Probability Distributions

This section quickly summarises some of the probability distributions used within this work. When analytically expressible, the maximum likelihood estimators of the parameters are given as a function of a random sample of size N: $\{x_k\}_{k\in[\![1;N]\!]}$ assumed to be compatible with the investigated distribution.

Please also note that there might exist some other equivalent parametrisation of these distributions. They have been omitted for the sake of simplicity. In particular, some distributions do not natively introduce any location parameter and could therefore be extended using a translation operator. For example, the exponential distribution should be generalised by extending its probability density function by equation (C.2.1).

$$E_{\lambda,\mu} : \left\{ \begin{array}{ccc} [\mu;\infty[& \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \lambda e^{-\lambda(x-\mu)} \end{array} \right. \tag{C.2.1}$$

C.2.1 Beta distribution

The Beta distribution has been widely used to model statistics with a finite support. Its probability density function is given in equation C.2.2.

$$\mathcal{B}_{a,b} : \left\{ \begin{array}{ccc} [y_1, y_2] & \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \frac{(x-y_1)^{a-1}(y_2-x)^{b-1}}{(y_2-y_1)^{a+b-1}B(a,b)} \end{array} \right. \tag{C.2.2}$$

The maximum likelihood estimator of the Beta distribution can be expressed by mean of the di-gamma function resulting in system (C.2.3) [Beckman and Tiet jen, 1978].

$$\left\{ \begin{array}{rcl} \Psi\left(\tilde{a}\right) - \Psi\left(\tilde{a}+\tilde{b}\right) & = & \ln\left(\prod_{k=1}^{N}\left(\frac{x_k-y_1}{y_2-y_1}\right)^{1/N}\right) = \ln G_1 \\ \Psi\left(\tilde{b}\right) - \Psi\left(\tilde{a}+\tilde{b}\right) & = & \ln\left(\prod_{k=1}^{N}\left(\frac{y_2-x_k}{y_2-y_1}\right)^{1/N}\right) = \ln G_2 \end{array} \right. \tag{C.2.3}$$

On the hypothesis that both shape parameters are strictly greater than one, the approximation of the bi-gamma function by the logarithm (see A.4.4) can be used to obtain an approximate solution.

$$\left\{ \begin{array}{rcl} \tilde{a}_0 & \approx & \frac{1}{2} + \frac{G_1}{2(1-G_1-G_2)} \\ \tilde{b}_0 & \approx & \frac{1}{2} + \frac{G_2}{2(1-G_1-G_2)} \end{array} \right. \tag{C.2.4}$$

In order to refine these solutions, a zero crossing of (C.2.3) is iteratively searched around $\left(\tilde{a}_0, \tilde{b}_0\right)$ and the solution is refined using dichotomy.

C.2.2 Cauchy Distribution

This distribution is also sometimes called Lorentz or Lorentzian distribution. It is fully defined through two parameters. A strictly positive real number, γ, also called *shape parameter*, and a real number, μ or *location parameter*, are enough to fully define the distribution.

$$\mathcal{C}_{\mu,\gamma} : \left\{ \begin{array}{ccc} \mathbb{R} & \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \frac{1}{\pi\gamma\left(1+\left(\frac{x-\mu}{\gamma}\right)^2\right)} \end{array} \right. \tag{C.2.5}$$

The Cauchy distribution can be interpreted as the distribution of a ratio of two independent normal random variables. In its general form, it has various applications in the field of physics such as describing shapes of spectral lines subjects to homogeneous broadening [Ott et al., 2013].

The maximum likelihood estimators of the parameters are not directly accessible but can be expressed through the system (C.2.6) which can be solved numerically [Ferguson, 1978].

$$\left\{ \begin{array}{rcl} \sum\limits_{k=1}^{N} \frac{x_k-\mu}{\gamma^2+(x_k-\mu)^2} & = & 0 \\ \\ \sum\limits_{k=1}^{N} \frac{\gamma^2}{\gamma^2+(x_k-\mu)^2} & = & 0 \end{array} \right. \tag{C.2.6}$$

C.2.3 Chi-Squared Distribution

This distribution is defined through one unique positive integer, k, called *degrees of freedom* of the χ^2 distribution. χ^2_k can indeed be understood as being the distribution of the sum of square values from k independent Gaussian distributions whose mean are 0 and variance equal to 1.

$k \in \mathbb{N}$

$$\chi^2_k : \left\{ \begin{array}{ccc} \mathbb{R}^+ & \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}} \end{array} \right. \tag{C.2.7}$$

The degree of freedom is here computed on its first moments: k being equal to the expectancy.

C.2.4 Gamma Distribution

The Gamma distribution is defined by two strictly positive real parameters, θ, the *scale parameter* and k, its *shape parameter*. The corresponding probability density function is given by equation (C.2.8). When k is an integer, the Gamma distribution is sometimes called the Erlang distribution. It can then be interpreted as representing the sum of k independent and identically distributed exponential random variables with rates equal to θ^{-1}.

$$\gamma_{k,\theta} : \left\{ \begin{array}{ccc} \mathbb{R}^+ & \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \frac{x^{k-1}}{\Gamma(k)\theta^k} e^{-\frac{x}{\theta}} \end{array} \right. \tag{C.2.8}$$

The maximum likelihood estimator leads to the system (C.2.9) [Harter and Moore, 1965].

$$\left\{ \begin{array}{rcl} \tilde{\theta}\tilde{k} & = & \frac{1}{N} \sum_{k=1}^{N} x_k \\ \ln\left(\tilde{\theta}\right) + \Psi\left(\tilde{k}\right) & = & \frac{1}{N} \sum_{k=1}^{N} \ln\left(x_k\right) \end{array} \right. \tag{C.2.9}$$

An iterative solver based on the the Newton-Raphson method has been proposed by [Choi and Wette, 1969]. Expressing $\tilde{\theta}$ as a function of $\tilde{k}$ yields equation (C.2.10), where Ψ is the di-gamma function.

$$\left\{ \begin{array}{rcl} \tilde{\theta} & = & \frac{1}{N\tilde{k}} \sum_{k=1}^{N} x_k \\ 0 & = & \ln\left(\tilde{k}\right) - \Psi\left(\tilde{k}\right) - \ln\left(\frac{1}{N} \sum_{k=1}^{N} x_k\right) + \frac{1}{N} \sum_{k=1}^{N} \ln x_k \\ & = & \ln\left(\tilde{k}\right) - \Psi\left(\tilde{k}\right) - M \end{array} \right. \tag{C.2.10}$$

The second term of equation (C.2.10) can then be solved iteratively by iterating on $\tilde{k}_n$:

$$\tilde{k}_{n+1} = \tilde{k}_n - \frac{\ln \tilde{k}_n - \Psi\left(\tilde{k}_n\right) - M}{\left(\tilde{k}_n\right)^{-1} - \Psi'\left(\tilde{k}_n\right)}$$

A possible initialisation can be based on the approximation of k proposed by [Thom, 1958]:

$$\tilde{k}_0 = \frac{1 + \sqrt{1 + \frac{4}{3}M}}{4M}$$

With

$$M = \ln\left(\frac{1}{N} \sum_{k=1}^{N} x_k\right) - \frac{1}{N} \sum_{k=1}^{N} \ln x_k$$

C.2.5 Exponential Distribution

This distribution is defined through a strictly positive real number, λ, usually called *rate parameter*. The exponential distribution can be interpreted as the time between two consecutive events of a Poisson process. It is commonly used to modell the life expectancy of electrical components. Its probability density function is given by equation (C.2.11).

$$E_\lambda : \left\{ \begin{array}{ccc} \mathbb{R}^+ & \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \lambda e^{-\lambda x} \end{array} \right. \tag{C.2.11}$$

The maximum likelihood estimator of the rate parameter is given by equation (C.2.12).

$$\tilde{\lambda} = \frac{N}{\sum\limits_{k=1}^{N} x_k} \tag{C.2.12}$$

C.2.6 Exponential-Power Distribution

The exponential power distribution, also called generalised normal distribution, is defined by three parameters:

- $\mu \in \mathbb{R}$, the *location parameter*.
- $\alpha \in \mathbb{R}^{+*}$, the *scale parameter*.
- $\beta \in \mathbb{R}^{+*}$, the *shape parameter*.

This is a generalisation of both the normal (i.e. $\beta = 2$) and the Laplace (i.e. $\beta = 1$) distributions. Its probability density function is given in equation (C.2.13).

$$Ep_{\mu,\alpha,\beta} : \left\{ \begin{array}{ccc} \mathbb{R} & \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \frac{\beta}{2\alpha}\Gamma\left(\frac{1}{\beta}\right)^{-1} e^{-\frac{|x-\mu|^\beta}{\alpha^\beta}} \end{array} \right. \tag{C.2.13}$$

The 'raw' maximum likelihood estimators of the generalised normal distribution are provided in equation (C.2.14), where Ψ is the di-gamma function.

$$\left\{ \begin{array}{rcl} \tilde{\mu} & = & \arg\min\limits_{\mu} \left(\sum\limits_{k=1}^{N} |x_k - \mu|^{\tilde{\beta}} \right) \\ \tilde{\alpha} & = & \left(\frac{\tilde{\beta}}{N} \sum\limits_{k=1}^{N} |x_k|^{\tilde{\beta}} \right)^{1/\tilde{\beta}} \\ 0 & = & 1 + \frac{\Psi\left(1/\tilde{\beta}\right)}{\tilde{\beta}} - \frac{\sum\limits_{k=1}^{N} |x_k|^{\tilde{\beta}} \ln\left(x_k\right)}{\sum\limits_{k=1}^{N} |x_k|^{\tilde{\beta}}} + \frac{\ln\left(\frac{\tilde{\beta}}{N} \sum\limits_{k=1}^{N} |x_k|^{\tilde{\beta}} \right)}{\tilde{\beta}} = g\left(\tilde{\beta}\right) \end{array} \right. \tag{C.2.14}$$

Since the equation $g\left(\tilde{\beta}\right) = 0$ is not directly invertible, [Do and Vetterli, 2002] proposes an iterative method based on the Newton-Raphson algorithm. Given an approximation of a root of g at step l, it can be updated using the recurrence equation (C.2.15).

$$\tilde{\beta}_{l+1} = \tilde{\beta}_l - \frac{g\left(\tilde{\beta}_l\right)}{g'\left(\tilde{\beta}_l\right)} \tag{C.2.15}$$

With g' being the first derivative of g: its expression is provided in (C.2.16) where Ψ' is the tri-gamma function as described in section A.4.8.

$$
\begin{aligned}
g'\left(\tilde{\beta}\right) \quad = \quad & \frac{1}{\tilde{\beta}^2} - \frac{\Psi\left(1/\tilde{\beta}\right)}{\tilde{\beta}^2} - \frac{\Psi'\left(1/\tilde{\beta}\right)}{\tilde{\beta}^3} - \frac{\sum\limits_{k=1}^{N} |x_k|^{\tilde{\beta}} \left(\ln\left(x_k\right)\right)^2}{\sum\limits_{k=1}^{N} |x_k|^{\tilde{\beta}}} + \left(\frac{\sum\limits_{k=1}^{N} |x_k|^{\tilde{\beta}} \ln |x_k|}{\sum\limits_{k=1}^{N} |x_k|^{\tilde{\beta}}} \right)^2 \\
& + \frac{\sum\limits_{k=1}^{N} |x_k|^{\tilde{\beta}} \ln |x_k|}{\tilde{\beta} \sum\limits_{k=1}^{N} |x_k|^{\tilde{\beta}}} - \frac{\ln\left(\tilde{\beta}/N \sum\limits_{k=1}^{N} |x_k|^{\tilde{\beta}}\right)}{\tilde{\beta}^2}
\end{aligned}
\tag{C.2.16}
$$

An initial guess of the shape parameter can be obtained using the inverse of the so-called generalised Gaussian ratio function (C.2.17) whose solution can be approximated by mean of a look-up table [Sharifi and Leon-Garcia, 1995; Do and Vetterli, 2002].

$$
r\left(\tilde{\beta}_0\right)^{-1} = \frac{\left(\Gamma\left(2/\tilde{\beta}_0\right)\right)^2}{\Gamma\left(1/\tilde{\beta}_0\right)\Gamma\left(3/\tilde{\beta}_0\right)} = \frac{\sum\limits_{k=1}^{N} |x_k|}{\sum\limits_{k=1}^{N} \left(x_k\right)^2}
\tag{C.2.17}
$$

Once the maximum likelihood estimator of the shape parameter has been retrieved, the location and scale parameters are straightforward to obtain.

C.2.7 Landau Distribution

The Landau distribution is a specific case of the levy alpha stable distribution. It can be expressed through equation (C.2.18), where r is any positive real number. It was initially introduced to model energy loses due to ionisation for charged particles going through thin layers of materials [Kölbig and Schorr, 1984]. Since it is not parameter dependant, this distribution has no maximum likelihood estimator.

$$
\mathcal{L} : \left\{ \begin{array}{ccc} \mathbb{R} & \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \frac{1}{2\pi i} \int_{r+i\infty}^{r+i\infty} e^{t \ln t + xt} dt \end{array} \right.
\tag{C.2.18}
$$

C.2.8 Laplace Distribution

This distribution is defined through a real parameter, μ or the *location parameter* and a purely positive real number a, its *scale parameter*. Its probability density function is given by equation (C.2.19). The distribution resulting from two independent and identically distributed exponential random variables is a Laplace distribution.

$$
L_{\mu,a} : \left\{ \begin{array}{ccc} \mathbb{R} & \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \frac{1}{2a} \exp\left(-\frac{-|x-\mu|}{a}\right) \end{array} \right.
\tag{C.2.19}
$$

The maximum likelihood for the location parameter, $\tilde{\mu}$, is simply the empirical median of the sample, while the estimator for the scale parameter is given by equation (C.2.20).

$$
\tilde{a} = \frac{1}{N} \sum_{k=1}^{N} |x_k - \tilde{\mu}|
\tag{C.2.20}
$$

C.2.9 Log-normal Distribution

The log-normal distribution is defined through two parameters. A strictly positive real number, σ or *shape parameter* and a real number, μ, called *log-scale parameter*. The corresponding probability density function is provided in equation (C.2.21).

$$\mathcal{LN}_{\mu,\sigma} : \begin{cases} \mathbb{R}^+ & \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} \end{cases} \tag{C.2.21}$$

The log-normal distribution can be interpreted as being the limit distribution corresponding to the multiplication of a large amount of independent (positive) random variables. In other words, this corresponds to the central limit theorem expressed through logarithms. The maximum likelihood of its parameters can be expressed analytically through system (C.2.22).

$$\begin{cases} \tilde{\mu} &= \frac{1}{N} \sum_{k=1}^{N} ln(x_k) \\ \tilde{\sigma} &= \sqrt{\frac{1}{N} \sum_{k=1}^{N} (ln(x_k) - \tilde{\mu})^2} \end{cases} \tag{C.2.22}$$

C.2.10 Normal Distribution

This is probably the best know and most widely used distribution. It is sometimes called the 'Gaussian' distribution. It is fully defined by its mean, μ, and its standard deviation, σ, which is of course a strictly positive real number. Its probability density function is given by equation (C.2.23). When $\mu = 0$ and $\sigma =$, this distribution is referred as the 'standard normal distribution'.

$$\mathcal{N}_{\mu,\sigma} : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{\sigma^2}} \end{cases} \tag{C.2.23}$$

Due to its relation to the central-limit theorem (see section 8.3.2), the normal distribution is widely used to model physical properties. It is moreover a stable distribution, i.e. the sum of two independent normal random variables follows a normal distribution too. It can therefore be seen as a special case of the α-stable distribution. Another extension of the normal distribution is the exponential power distribution.

The maximum likelihood estimators of its parameters are given in by the system (C.2.24).

$$\begin{cases} \tilde{\mu} &= \sum_{k=1}^{N} x_k \\ \tilde{\sigma} &= \sqrt{\sum_{k=1}^{N} (x_k - \tilde{\mu})^2} \end{cases} \tag{C.2.24}$$

C.2.11 Rayleigh Distribution

The Raleigh distribution is characterised by a strictly positive real number, σ, called *scale parameter*. This distribution can be used to model the norm of a vector, whose components are following independent and identically distributed normal distributions. Its probability density function can be expressed by equation (C.2.25).

$$\mathcal{R}_\sigma : \begin{cases} \mathbb{R}^+ & \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \frac{x}{\sigma^2} e^{-\frac{x^2}{\sigma^2}} \end{cases} \tag{C.2.25}$$

The maximum likelihood estimator of the scale parameter is given by equation (C.2.26).

$$\tilde{\sigma} = \sqrt{\frac{1}{2N} \sum_{k=1}^{N} x_k^2} \tag{C.2.26}$$

C.2.12 Student's t-Distribution

The t-distribution is fully defined through a single parameter: a strictly positive real number, ν or *degree of freedom*.

$$\mathcal{T}_\nu : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R}^+ \\ x & \rightarrow & \frac{\Gamma(0.5(\nu+1))}{\sqrt{\nu\pi}\,\Gamma(0.5\nu)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \end{cases} \tag{C.2.27}$$

C.3 About Stable Distribution

In the following, exclusive use of the S_0 parametrisation as explained in equation 8.4.4 is made.

C.3.1 Special Cases

Cauchy Distribution

Property 2. $\forall\,(\gamma,\delta)\in(\mathbb{R}^{+\star}\times\mathbb{R}),\ \forall x\in\mathbb{R},$

$$\mathcal{M}_{1,0,\gamma,\delta}\,(x)=\mathcal{C}_{\delta/2\pi,\gamma/2\pi}(x) \tag{C.3.1}$$

Proof. The complete proof is somewhat long and makes uses of notions about holomorphic functions (e.g. the Cauchy integration theorem) which would be too long to introduce here. We therefore admit that the characteristic function of the Cauchy distribution is given by:

$$\check{\mathcal{C}}_{\mu,\gamma}(\omega)=e^{2i\pi\omega\mu-2\pi\gamma|\omega|}$$

Setting $\alpha=1$ and $\beta=0$ in 8.4.4 directly gives the correspondence between both characteristic functions, with $\delta=\mu/(2\pi)$ and $\gamma=\alpha/(2\pi)$ □

Normal Distribution

As stated in section 8.4.1, the alpha stable distribution with α set to two is just a Gaussian distribution [Nolan, 2013].

Property 3. $\forall\,(\beta,\gamma,\delta)\in([-1;1]\times\mathbb{R}^{+\star}\times\mathbb{R}),\ \forall x\in\mathbb{R},$

$$\mathcal{M}_{2,\beta,\gamma,\delta}\,(x)=\mathcal{N}_{\delta/2\pi,\gamma/\sqrt{2}\pi}(x) \tag{C.3.2}$$

Proof. To start with, the β term of the characteristic function can be ignored since $\tan\alpha\pi/2=\tan\pi=0$. In other words, the characteristic function becomes

$$\widetilde{\mathcal{M}}_{2,\beta,\gamma,\delta}\,(\omega)=e^{i\delta\omega-\gamma^2\omega^2}$$

To obtain the probability density function from the characteristic function, one has to apply a Fourier transformation. The term $e^{-\gamma^2\omega^2}$ is a Gaussian-shaped function, whose Fourier transform is $\frac{\sqrt{\pi}}{\gamma}e^{-\pi^2x^2/\gamma^2}$ (see section D.1.2, with $a=\gamma^2$). The phase introduced by $e^{i\delta\omega}$ only induces an offset of $\delta/2\pi$ on the final result (see section D.2). This yields the following probability density function:

$$\mathcal{M}_{2,\beta,\gamma,\delta}=\frac{\sqrt{\pi}}{\gamma}e^{-\pi^2\frac{(x-\delta/2\pi)^2}{\gamma^2}}$$

By setting $\mu=\delta/(2\pi)$ and $\sigma=\gamma/\left(\sqrt{2}\pi\right)$ this can be rewritten in a more obviously Gaussian form:

$$\mathcal{M}_{2,\beta,\gamma,\delta}=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}=\mathcal{N}_{\mu,\sigma}(x)$$

□

C.3.2 Parameter Estimations

This section presents some additional information to the details provided in section 8.4.4. The following paragraphs summarise the parameter estimations methods along with their main pros and cons.

Quantile-Based Methods

The following methods base their parameter estimations on evaluations of quantiles, in the context of the S_1 parametrisation. In the following, Q_p denotes the p^{th} quantile. Given N samples $\{X_k\}_{k\in[\![1;N]\!]}$, sorted in increasing order: Q_p can be approximated by X_{pN}. When pN is not an integer, Q_f can be approximated by a linear interpolation between the pN floor (i.e $\lfloor pN \rfloor$) and ceiling (i.e. $\lceil pN \rceil$) :

$$Q_f = (pN - \lfloor pN \rfloor)\, X_{\lfloor pN \rfloor} + (\lceil pN \rceil - pN)\, X_{\lceil pN \rceil}$$

The basics of these methods started with the seminal work of Fama and Roll: [Fama and Roll, 1968, 1971]. They developed a method to estimate the stability and shape parameters, provided that the skewness parameter, β, is zero, based on simple quantile based estimators. The scale parameter is first estimated by

$$\gamma = \frac{Q_{0.72} - Q_{0.28}}{1.654}$$

The stability parameter could then be estimated by comparing, for p typically ranging around 0.96 ± 0.01, the ratio E_α with tabulated values computed with different values of α.

$$E_\alpha = \frac{Q_p - Q_{1-p}}{2\gamma}$$

This method however suffers from several drawback. To start with, it is restricted to the symmetrical case (i.e. $\beta = 0$) and when the stability parameter is greater than one ($\alpha \in [1; 2]$). It furthermore presents a slight bias on the estimators of the stability and scale parameters.

An improvement of the algorithm was presented in [McCulloch, 1986]. It, in particular, discards estimation biases and allows all parameter ranges, except for the stability parameter which has to be greater than 0.5. His methods start by estimating α and β (see equations C.3.3 C.3.4) which are only dependant on the stability and skewness parameters. Interpolating over look-up tables yields their estimations.

$$\nu_\alpha = \frac{Q_{0.95} - Q_{0.05}}{Q_{0.75} - Q_{0.25}} \tag{C.3.3}$$

$$\nu_\beta = \frac{Q_{0.95} + Q_{0.05} - 2Q_{0.5}}{Q_{0.95} - Q_{0.05}} \tag{C.3.4}$$

Once α and β have been computed it is possible to retrieve an estimation of the scale parameter using a function of α, β: ϕ, whose values are tabulated.

$$\gamma = \frac{Q_{0.75} - Q_{0.25}}{\phi\,(\alpha, \beta)} \tag{C.3.5}$$

The estimation of the location parameter is limited by its discontinuity around $\alpha = 1$. This issue is addressed by switching implicitly to the $\mathcal{M}$ parametrisation and solving equation C.3.6, where ϕ_2 is obtained through another look-up table as a function of α and β.

$$\delta = Q_{0.5} + \gamma\phi_2\,(\alpha, \beta) - \beta\gamma \tan(\pi\alpha/2) \tag{C.3.6}$$

The McCulloch method being mostly based on interpolation over look-up tables, it is in practice very fast. On the other hand, its precision is comparably low [Bates and McLaughlin, 1997; Čížek et al., 2005; McCulloch, 1986; Nolan, 2013], especially for α close to one.

Maximum Likelihood Estimation

The main issue in performing any maximum likelihood estimation of the parameters of a stable distribution is its lack of analytical expression. This difficulty can be by-passed in two ways. An estimation of the actual PDF can be retrieved from the characteristic function by means of Fast Fourier Transforms (FFT) [Mittnik et al., 1999a,b] or by means of a direct computation [Čížek et al., 2005; Nolan, 2013]. Both methods are comparable in terms of efficiency: they provide very accurate results at the cost of an extremely long computational time (about four orders of magnitude with respect to the [Kogon and Williams, 1995] method [Čížek et al., 2005].

Although slightly more precise than the characteristic function methods, the main drawback of the maximum likelihood method is their huge complexity which only yields negligible improvements of the results.

C.3.3 Continuity of Parameters

This section reviews the limit cases of the parameters in the case of the S_0 parametrisation, when $\alpha \to 1$ and in particular for the β-term:

$$\begin{aligned}
\tan(\pi\alpha/2)\left(x^{1-\alpha}-1\right) &= \tfrac{\sin(\pi\alpha/2)}{\cos(\pi\alpha/2)}\left(e^{(1-\alpha)\ln(x)}-1\right) \\
&= \tfrac{\cos(\pi/2-\pi\alpha/2)}{\sin(\pi/2-\pi\alpha/2)}\left(e^{(1-\alpha)\ln(x)}-1\right) \\
&= \left(\tfrac{1}{(1-\alpha)\pi/2}+o\left(1-\alpha\right)\right)\left(1+(1-\alpha)\ln(x)+o\left((1-\alpha)^2\right)-1\right) \\
&\sim \tfrac{2}{\pi}\ln(x)
\end{aligned} \tag{C.3.7}$$

therefore the limit of the β-term in $\mathcal{M}_{\alpha\neq 1,\beta,\gamma,\delta}$ when $\alpha \to 1$ exists and is equal to

$$\begin{aligned}
\lim_{\alpha\to 1}\left[i\beta sgn(\omega)\tan\left(\pi\alpha/2\right)\left(|\gamma\omega|^{1-\alpha}\right)\right] &= i\beta sgn(\omega)\tfrac{2}{\pi}\log|\gamma\omega| \\
&= i\tfrac{2}{\pi}\beta sgn(\omega)\left(\log\gamma+\log|\omega|\right)
\end{aligned} \tag{C.3.8}$$

The continuity for the other parameters and/or for the other terms is straightforward.

Appendix D

Fourier Transform

> Mathematical analysis is as extensive as nature itself; it defines all perceptible relations, measures times, spaces, forces, temperatures ; this difficult science is formed slowly, but it preserves every principle which it has once acquired; it grows and strengthens itself incessantly in the midst of the many variations and errors of the human mind.
>
> Joseph Fourier - Scientist

D.1 Principles & Definitions

D.1.1 Definition

In the following, a function f: $\mathbb{R} \to \mathbb{C}$ is considered. Its Fourier transform is written $\hat{f}$ (see equation D.1.1), while its inverse Fourier transform is noted $\check{f}$ (see equation D.1.2).

$$\hat{f}\begin{cases} \mathbb{R} & \to & \mathbb{C} \\ \omega & \to & \int\limits_{-\infty}^{\infty} f(t)e^{-2i\pi\omega t}dt \end{cases} \tag{D.1.1}$$

$$\check{f}\begin{cases} \mathbb{R} & \to & \mathbb{C} \\ t & \to & \int\limits_{-\infty}^{\infty} f(\omega)e^{2i\pi\omega t}d\omega \end{cases} \tag{D.1.2}$$

The Fourier transform of a function taken at ω can be understood as the function projection on the ω frequency. One can also note that the Fourier transform is not defined for every function: it is however well defined for functions from $\mathbb{L}^1(\mathbb{R}) \cap \mathbb{L}^2(\mathbb{R})$ which correspond to the vast majority of physics-based functions.

Similarly to the Fourier transform, the Mellin transform of f - written $\mathcal{F}$ - , can be defined in terms of the Fourier transform of the composition $Fo\exp$:

$$\mathcal{F}\begin{cases} \mathbb{R} & \to & \mathbb{C} \\ \omega & \to & \widehat{f(\exp^x)}(\omega) = \int\limits_{0}^{\infty} f(x)x^{\omega-1}dx \end{cases} \tag{D.1.3}$$

D.1.2 Example: Gaussian

Property 4. *The Fourier transform of a Gaussian-shaped function is another Gaussian shaped function. For all strictly positive constant a, if $f : x \in \mathbb{R} \rightarrow e^{-ax^2}$, then*

$$\hat{f}(\omega) = \sqrt{\frac{\pi}{a}} e^{-\pi^2\omega^2/a} \tag{D.1.4}$$

Proof. f is derivable and $f'(x) = -2axf(x)$. Applying a Fourier transform to both side of the equation leads to $\hat{f}'(\omega) = 2i\pi\omega$ and $\hat{f}'(\omega) = 2a\hat{f}'(\omega)/(2i\pi)$. Combining the two equations leads to a first order differential equation in $\hat{f}$, $\hat{f}(\omega) + 2\pi^2/a\omega\hat{f}(\omega) = 0$ which admits a unique solution:

$$\hat{f}(\omega) = \hat{f}(0)e^{-\pi^2\omega^2/a}$$

The constant coefficient $\hat{f}(0)$ is straightforward to obtain by using the well known Euler-Poisson integral:

$$\hat{f}(0) = \int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$$

This yields the final result:

$$\hat{f}(\omega) = \sqrt{\frac{\pi}{a}} e^{-\pi^2\omega^2/a}$$

□

D.1.3 Fast Fourier Transform

One of the main advantages of the Fourier transform resides in its potential for efficient numerical implementations, especially for sample sizes which are power of two. The main idea behind this class is to divide a discrete Fourier transform into two sub-problems of equal sizes. The corresponding algorithms - called FFT (or Fast Fourier Transform) have a complexity in $\mathcal{O}(N\log(N))$, N being the number of sample points in the function to be transformed.

D.2 Properties

The practical interest of the Fourier transform lies in its many interesting properties.

To start with, the Fourier transform is actually reversible: $f = \hat{\hat{f}} = \check{\check{f}}$. A further advantage of the Fourier transform is its linearity: given two scalars a and b as well as two functions f and g, $\widehat{af + bg} = a\hat{f} + b\hat{g}$, thus allowing straightforward computations in most of the cases.

Furthermore, the Fourier transform is quite unaffected by scaling and translation, given a scalar a and two functions f and g:

- If $g(x) = f(x - a)$ then $\hat{g}(\omega) = \exp(-2i\pi\omega a).\hat{f}(\omega)$.
- If $g(x) = f(ax)$ then $\hat{g}(\omega) = \hat{f}(\omega/a)/|a|$.

Finally, the Fourier transform of the convolution of two terms is the product of both Fourier transforms: if $h(x) = (f \star g)(x)$, then

$$\hat{h} = \hat{f} \cdot \hat{g} \tag{D.2.1}$$

Appendix E

Linear Systems

You must always invert!

Carl G.J. Jacobi - Mathematician

E.1 Foreword

This appendix explicates some of the notions related to solving the system $A \cdot X = B$ for X, when A is a known N by N matrix and B a known vector. Again, the focus is set on the underlying ideas and concepts, technical mathematical subtleties being left aside.

In the following, $\mathcal{M}_N(\mathbb{R})$ (resp. $\mathcal{M}_N(\mathbb{C})$) denotes the set of square matrices of order N with values in $\mathbb{R}$ (resp. $\mathbb{C}$). It should be noted that $\mathbb{R}^N \subset \mathbb{C}^N$. As a consequence, and since in some cases the use of $\mathbb{C}$ is required, definitions and examples are only made on $\mathbb{C}$.

E.2 Conditioning

Before going further in the topic, a few definitions have to be provided

Definition E.2.1 (Surrogate Matricial Norm)**.** *Given a vectorial norm, $\|\bullet\|$ defined on $\mathbb{C}^N$, its surrogate matricial norm in $\mathcal{M}_N(\mathbb{C})$ is defined for a matrix A by equation E.2.1:*

$$\|A\| = \sup_{X \in \mathbb{C}^N, X \neq 0} \frac{\|A \cdot X\|}{\|X\|} \tag{E.2.1}$$

Surrogate matricial norms are usually defined by means on one of the three classical norms on $\mathbb{C}^N$: given a vector $X = (x_1, x_2, \dots x_N)$:

- $\|X\|_1 = \sum_{k=1}^{N} |x_k|$
- $\|X\|_2 = \sqrt{\sum_{k=1}^{N} |x_k|^2}$
- $\|X\|_\infty = \max_{k \in [\![1;N]\!]} |x_k|$

The corresponding surrogate norms on matrix is, by extension, noted similarly (respectively $\|A\|_1$, $\|A\|_2$ and $\|A\|_\infty$) and called matricial norms.

Once a matricial norm has been defined, it is possible to express its condition number: $cond\,(A)$ as in definition E.2.2.

Definition E.2.2. *[Condition Number] Given a invertible matrix A from $\mathcal{M}_N(\mathbb{C})$ and a surrogate norm $\|\bullet\|$, the condition number is defined as:*

$$cond\,(A) = \|A\| \cdot \|A^{-1}\| \tag{E.2.2}$$

It can then be shown [Allaire, 2005], that if X_0 and $X_0 + \Delta X$ are respectively solutions of the systems $A \cdot X = B$ and $A \cdot X = B + \Delta B$ then equation E.2.3 is valid.

$$\frac{\|\Delta X_0\|}{\|X_0\|} \leqslant cond(A) \frac{\|\Delta B\|}{\|B\|} \tag{E.2.3}$$

Similarly if X_0 and $X_0 + \Delta X$ are respectively solutions of the systems $A \cdot X = B$ and $(A + \Delta A) \cdot (X + \Delta X) = B$ then equation E.2.4 is valid too.

$$\frac{\|\Delta X_0\|}{\|X_0 + \Delta X_0\|} \leqslant cond(A) \frac{\|\Delta A\|}{\|A\|} \tag{E.2.4}$$

It results that, when A has a high condition number, small uncertainties and approximations in A or B can potentially lead to high changes in X, thus limiting or preventing a solver to converge properly. It should furthermore be added that since solvers are typically numerical, they are working on approximations of real numbers, thus leading in any case to small rounding errors in both A and B.

When A has a bad conditioning, it is sometimes possible to lower its condition number by means of a pre-conditioning invertible matrix P and solving the equivalent system $P^{-1}AX = P^{-1}B$. The choice of the right P is however a difficult task.

E.3 LU-Decomposition

Although not the only numerical method for solving a linear system, the LU decomposition relies on few requirements to converge and is therefore widely used. For matrices that have more properties (symmetry, diagonally dominant) some faster methods can be used.

The idea behind the LU decomposition is to decompose A as the product of a lower triangular matrix (L) and an upper triangular matrix (U) (equation E.3.1) where the matrix coefficient of U and L can be iteratively computed by means of equation E.3.2.

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ l_{2,1} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n,1} & l_{n,2} & \dots & 1 \end{bmatrix} \cdot \begin{bmatrix} u_{1,1} & u_{1,2} & \dots & u_{1,n} \\ 0 & u_{2,2} & \dots & u_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{n,n} \end{bmatrix} \tag{E.3.1}$$

$$\begin{cases} u_{k,p} = a_{k,p} - \sum_{m=1}^{k-1} l_{k,m} u_{m,p} & \text{for } k \leqslant l \\ l_{k,p} = \dfrac{a_{k,p} - \sum_{m=1}^{p-1} l_{k,m} u_{m,p}}{u_{p,p}} & \text{for } p+1 \leqslant k \end{cases} \tag{E.3.2}$$

Once U and L have been computed, solving $A \cdot X = B$ can be simply decomposed into the successive solving of two diagonal systems: $LY = B$ and $UX = Y$. The solver itself requires $\mathcal{O}\left(n^2\right)$ operations. However, it needs to have already performed the decomposition of A into L and U as in equation E.3.2 which has a complexity in $\mathcal{O}\left(n^3\right)$.

Appendix F

Data-Sets Used

If you torture the data enough,
Nature will always confess.

Ronald Coase - Economist

F.1 Kaufbeuren 2012

The Kaufbeuren site, a former military airfield, is used as a calibration and test site for airborne sensors by the German Aerospace Center. In 2012, the opportunity to propose a specific set-up over the area was given, thus allowing to obtain a 'tuned' data-set. The retained set-up consisted of a total of 10 flight-lines:

- Three flight-lines, $4A$, $4B$ and $4C$ were acquired using the same acquisition set-up, hence leading to three overlapping flight-lines.
- Eight flight-lines, 1, 2, 3, $4A$, $4B$, $4C$, 5 and 6 were acquired with parallel orientations and fringes overlaps, hence being similar to a real case survey.
- Two flight-lines, 7 and 8 were acquired with different orientations, hence allowing a better representativeness of the view-angle changes.

The test site corresponds to a relatively flat area with a broad variety of land covers including large concrete areas (e.g. runway), forested areas, short vegetation areas (e.g. fields) as well as urban structures and some geometric targets (e.g. solar panel arrays). The acquisition took place on the first of August 2012 around eleven in the morning (local time).

The data-set itself was acquired by means of the HySpex sensor, and is therefore divided into two sub-images, the first containing the visible and near infra-red part of the imagery, the second one including the short-wave infra-red of the spectra. Since both were acquired with different swath and ground resolutions, they were dealt with as if they were two separate data-sets. A quick look for the visible and near infra-red part is presented on figure F.1 (Red: band #75 at 684 nm - Green: band #45 at 575 nm - Blue: band #10 at 448 nm).

A subset of the flight-line $4A$ was used to perform experiments described in chapter 4. The corresponding imagery is displayed on figure F.2 (Red: band #75 at 684 nm - Green: band #45 at 575 nm - Blue: band #10 at 448 nm).

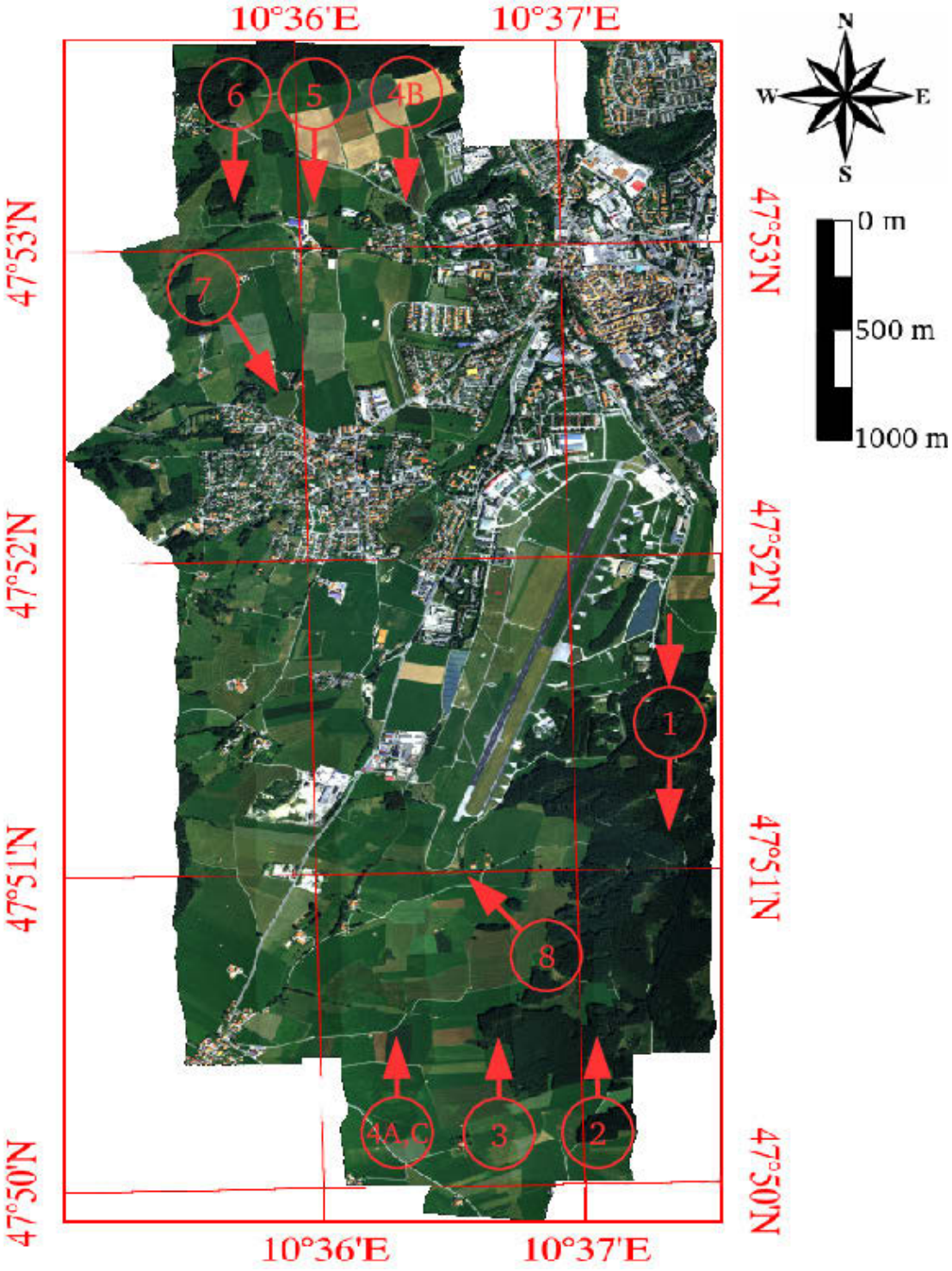

Figure F.1: Quicklook of the Kaufbeuren test site (Germany) - Sensor: HySpex, GSD for VIS-NIR 0.75m, for SWIR: 1.5m - arrowed numbers indicate the flight-line positions and heading.

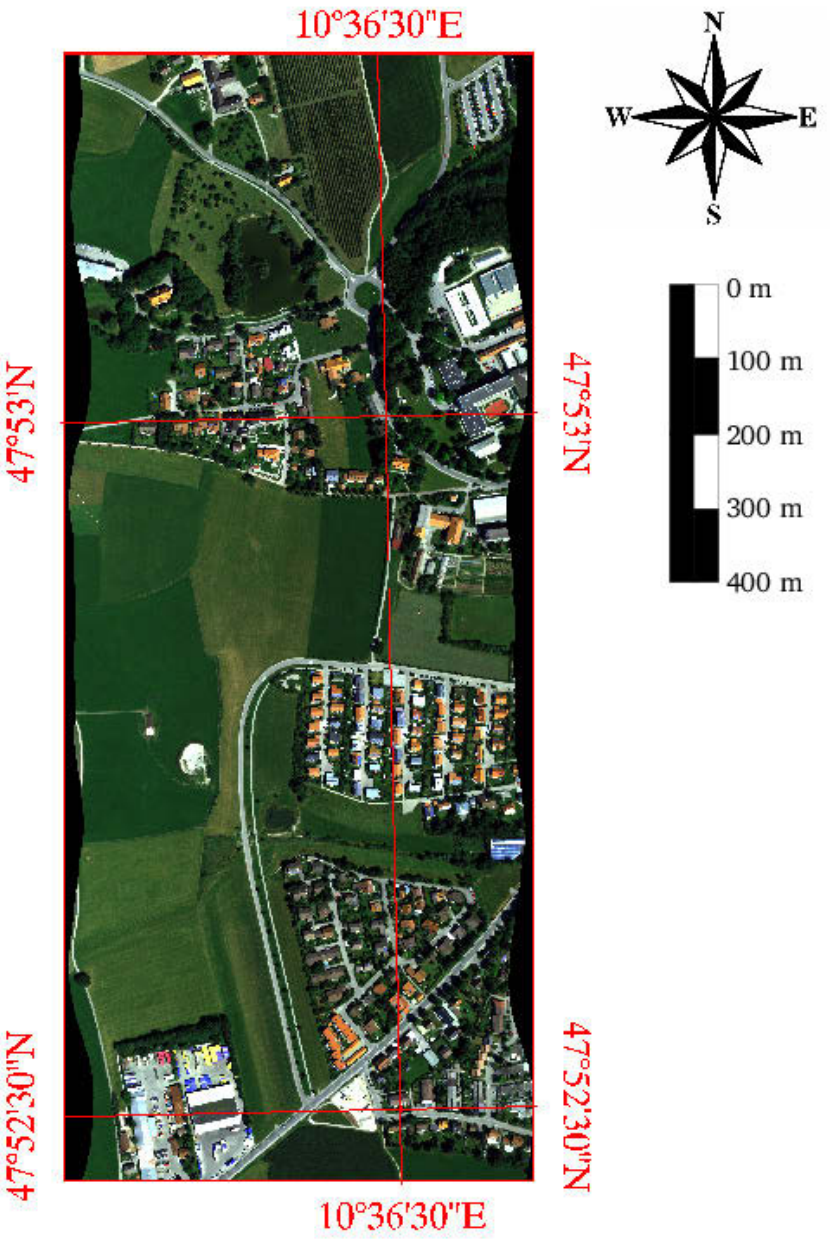

Figure F.2: Subset of flight-line 4*A* from the Kaufbeuren scene used in chapter 4.

F.2 Sokolov

The Sokolov test site corresponds to a lignite open-pit mining area located around the town of Sokolov in the Czech Republic. The scene corresponds to a hilly area, with some strong relief contrasts within the open-pits, making the retrieval of an accurate digital elevation model difficult. This does in turn lead to quite large errors during the ortho-rectification step. As for the land-cover, this scene also presents a large variability. The open-pits and dump sites exhibit bare soils profiles, making mineral mapping possible. Remaining areas consist of of small forests, open fields, water bodies (mostly former pits, rehabilitated into lakes), cultivated fields as well as some urban areas. As for the Kaufbeuren site, the Sokolov site was investigated during many field campaigns ranging from 2006 to 2011 (FP-7 project EO-Miners: [EO-Miners, 2014]).

F.2.1 HyMap - 2010

The first data-set to be used was acquired with the HyMap sensor (see section G.3) on the 21^{th} of August 2010 around noon (local time). The data includes seven parallel flight-lines, acquired with a ground resolution of four meters. The oblique North-East to South-West acquisition pattern was chosen in order to place the flight-lines parallel to the solar-plane, thus minimising the changes due to relative orientation with respect to the sun. A quicklook is provided on figure F.3 (red: band #15 at 666 nm - green: band #10 at 589 nm - blue: band #3 at 483 nm).

F.2.2 AHS - 2011

Beside classical hyperspectral data in the optical domain, a further data-set was acquired in the thermal domain. This was achieved by means of the AHS multispectral scanner (see section G.2). The corresponding data-set was acquired during a night acquisition on the 22^{nd} of July 2011, around $23:30$ local time.

A quicklook of the imagery is displayed on figure F.4 (at-sensor radiance: band #75 at 10143 nm, pixel size: 5 m). Please note that, since the imagery was made in the thermal domain, displayed data does not correspond to ground reflectance but only to 'at sensor radiance'. For more information about the pre-processing of this specific data-set, please refer to [Fischer and Ehrler, 2012; Ehrler, 2014]. The data-set itself is made of nine parallel flight-lines following a north-south axis, with one cross-line following an east-west axis.

F.2.3 HySpex - 2013

A final data-set was acquired over Sokolov by means of the HySpex sensor on the second of August 2013 around $13:00$ (local time). As for every HySpex data-set, the imagery is actually divided into two parts. The first corresponds to the visible to near infra-red part of the spectra and was acquired with a ground resolution of about 2.40 meters. The second corresponds to the short-wave infra-red detector and corresponds to a ground resolution of about 4.80 meters. The data-set consists of a total of 17 flight-lines, two of them having been flown on an east-west axis, the remaining ones having been acquired over a north-south axis. A quicklook is provided on figure F.5 (Red: band #75 at 684 nm - Green: band #45 at 575 nm - Blue: band #10 at 448 nm).

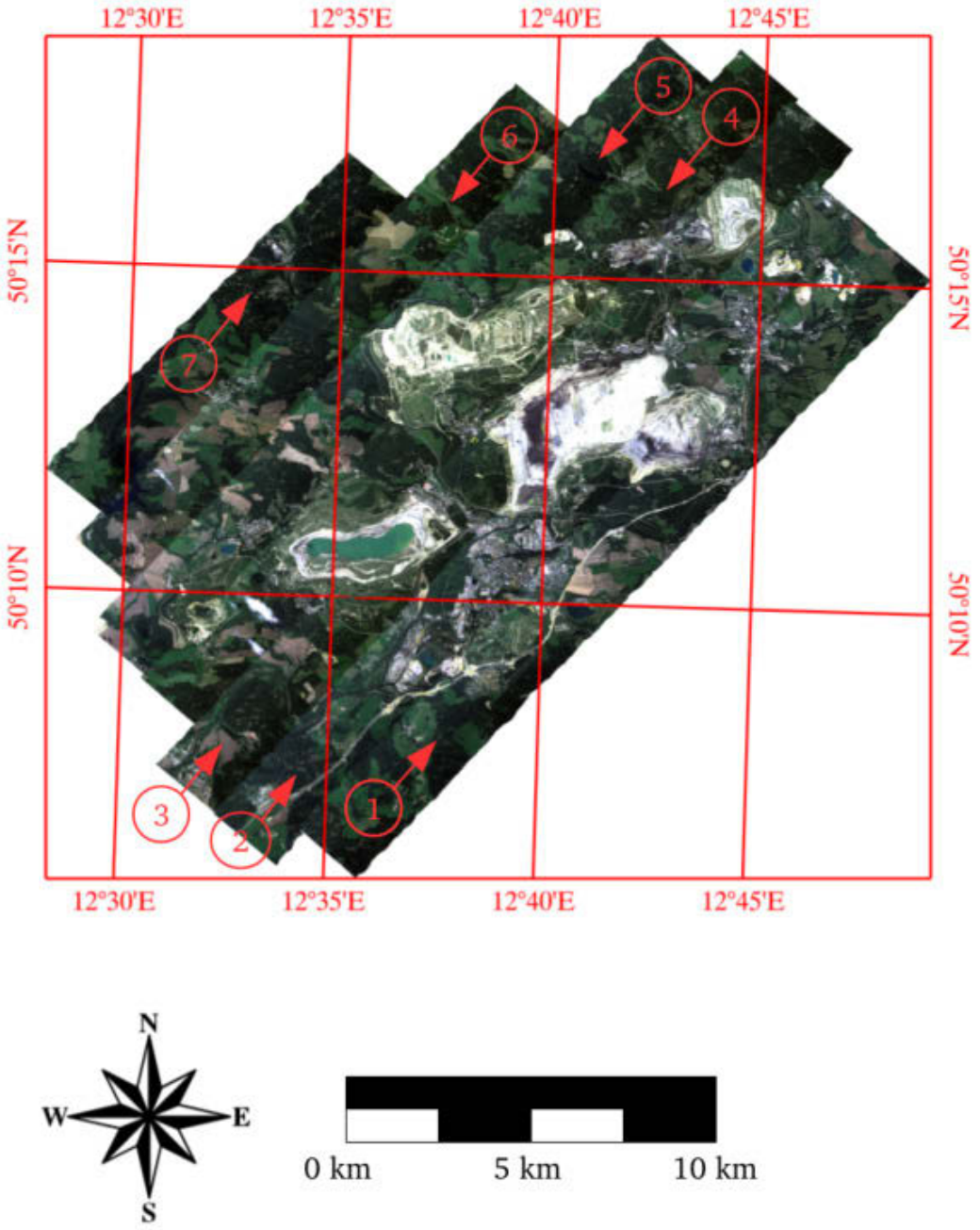

Figure F.3: Quicklook of the HyMap - Sokolov 2010 data-set. Arrowed numbers indicate flight-line as well as their flight direction.

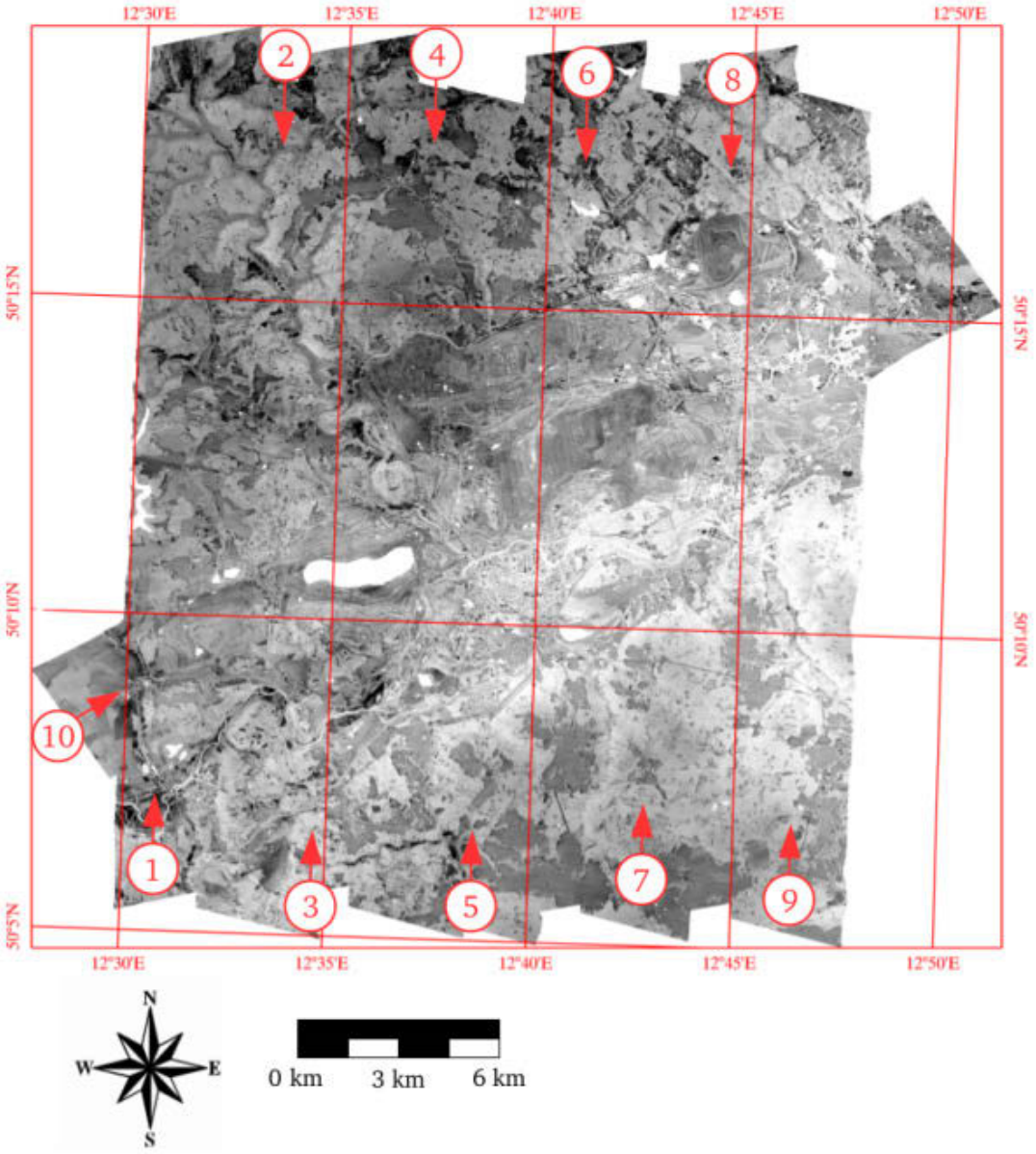

Figure F.4: Quicklook of the AHS - Sokolov 2011 data-set. Arrowed numbers indicate flight-line as well as their flight direction.

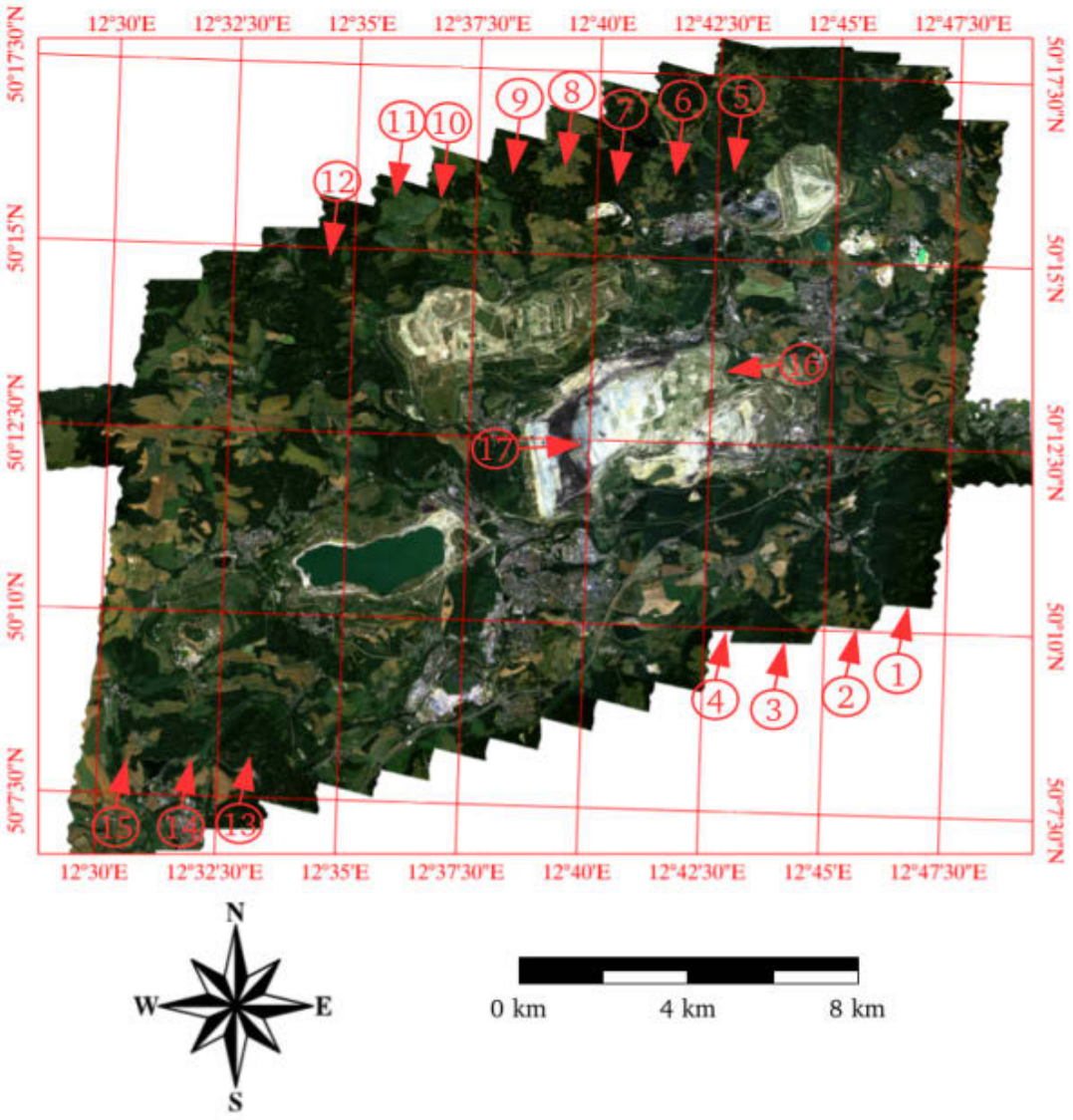

Figure F.5: Quicklook of the HySpex - Sokolov 2013 data-set. Arrowed numbers indicate flight-line as well as their flight direction.

F.3 Witbank

The Witbank test site is located in the Mpumalanga province of the Republic of South Africa. The scene corresponds to a mixture of formal and informal settlements, industrial areas (smelter), water bodies (river, reservoirs), raw material dumps and semi-arid grass land. Due to former underground coal mining activities subsidence is to be observed in many areas, thus resulting in a complex terrain geometry. Several other environmental issues, such as underground coal fires and acid mine drainage can be observed.

This work is based on two data-sets acquired over this test site: a satellite stereo-view acquisition described in section F.3.1 and an airborne hyperspectral data-set described in section F.3.2.

F.3.1 World-View-II - 2011

This data-set is made of a stereo-pair of multi-spectral spaceborne imagery acquired by means of the World-View-II satellite (see appendix G.5). Imagery was acquired on the 7^{th} of December 2011 around $10:40$ (local time). Both acquisitions were distant of about one and a half minutes on a descending orbit. A snapshot of the second image is provided on figure F.6 - Red: Band #2 at 659 nm, Green: Band #3 at 546 nm, Blue: Band #4 at 478 nm.

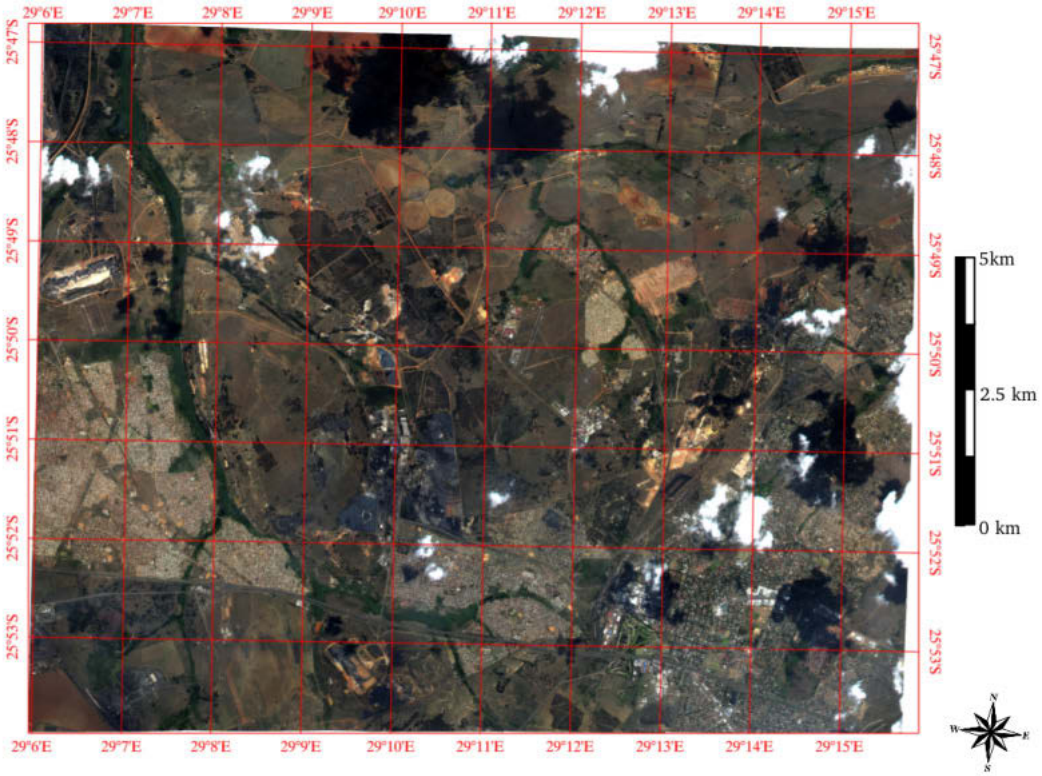

Figure F.6: Quicklook of the World-View II multi-spectral stereo-pair acquired over Witbank (2011): second acquisition.

F.3.2 Aisa Dual - 2012

The AISA Dual data-set acquired over Witbank is quite complex. The data was not acquired on a single date but over three days: October the 19^{th} and November the 12^{th} and 13^{th}, 2012, acquisition times ranging from $9:00$ to $15:00$ totalling 68 flight-lines. An overview is provided on figure F.7 - Red: band # 50 at 621 nm, Green: band #34 at 547 nm, Blue: band #13 at 450 nm. For the sake of readability, the flight-line numbering was not displayed on the figure. Line numbering is made based on longitude: 1 is the western-most line, 68 being the eastern most. Flight-lines 1 to 27, 37 to 40, 64 and 68 were acquired in October, lines 28, 29, 41 to 43, 53 and 63 were acquired on the 12^{th} of November. The remaining ones were acquired on November the 13^{th}. Flight-lines 1 to 12, 37 to 42, 44 to 56, 64 to 65 and 67 were flown from south to north. The remaining ones were flown from north to south.

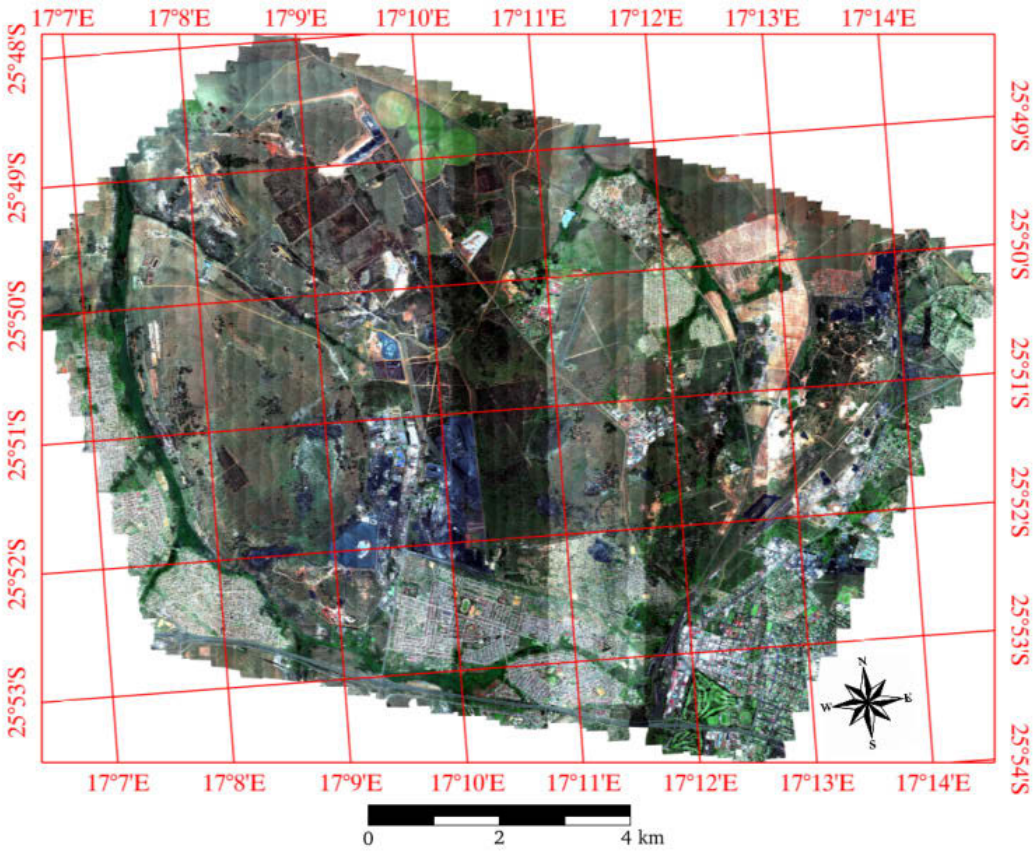

Figure F.7: Quicklook of the AISA-Dual - Witbank 2012 data-set.

Appendix G

Sensors

> Pictures produced by camera can resemble paintings or drawings in presenting recognizable images of physical objects. But they have also characteristics of their own, of which the following two are relevant here: first the photograph acquires some of its unique visual properties through the technique of mechanical recording; and second, it supplies the viewer with a specific kind of experience, which depends on his being aware of the picture's mechanical origin. To put it more simply: (1) the picture is coproduced by nature and man and in some ways looks strikingly like nature, and (2) the picture is viewed as something being by nature.
>
> Rudolf Arnheim - Film Theorist

This appendix summarises the main characteristics of the different sensor systems used within this work. They are listed by alphabetical order and their characteristics are given under nominal conditions of operations.

G.1 AISA Dual

The AISA Dual sensor was used to acquire a single data-set over Witbank: see section F.3.2. Its characteristics are given in table G.1 [Specim, 2012].

AISA Dual	
Number of Bands	360
Spectral Coverage	396 to 2454 nm
Spectral sampling	2.3 to 23.2 nm
Bands FWHM	4.4 to 10.7 nm
FOV	24°
Pixel size	1x1 m^2

Table G.1: AISA Dual sensor data-sheet (nominal operations).

G.2 AHS

The AHS sensor system was used to acquire a single data-set over Sokolov (see section F.2.2). Since only the 4^{th} port of the sensor - corresponding to the long-wave infra-red - was used, it is the only one described on table G.2 [Fernández-Renau et al., 2005].

AHS (4^{th} port)	
Number of Bands	10
Spectral Coverage	8310 to 12952 nm
Spectral sampling	443 to 606 nm
Bands FWHM	412 to 566 nm
FOV	90°
Pixel size	2.5x2.5 m^2

Table G.2: AHS sensor data-sheet (nominal operations).

G.3 HyMap

The HyMap sensor system was the main sensor-system to be used by DLR until the acquisition of the new generation HySpex. Its characteristics are summarised in table G.3 [Cocks et al., 1998].

HyMap	
Number of Bands	128
Spectral Coverage	454 to 2485 nm
Spectral sampling	13 to 17 nm
Bands FWHM	15 to 20 nm
FOV	61.3°
Pixel size	4x4 m^2

Table G.3: HyMap sensor data-sheet (nominal operations).

G.4 HySpex as operated by DLR

The current sensor system operated by DLR is made of two HySpex The HySpex sensor operated by DLR is actually made of two HySpex cameras [Norsk Elektro Optikk (NEO), 2009; Baumgartner et al., 2012]: namely a $VNIR$-1600 - imaging spectra in the visible an near infra-red - in conjunction with a $SWIR$-320m-e for the short wave infra-red. For the sake of simplicity, the complete system is however only referred to as HySpex. The system can furthermore be operated with a field of view expander which allows a multiplication for the FOV by a factor two. Details of the sensors are given in table G.4.

HySpex VNIR-1600	
Number of Bands	160
Spectral Coverage	456 to 992 nm
Spectral sampling	3.7 nm
Bands FWHM	3.5 nm
FOV	17°
Pixel size	0.75x0.75 m^2

HySpex SWIR-320m-e	
Number of Bands	256
Spectral Coverage	1000 to 2497 nm
Spectral sampling	6 nm
Bands FWHM	5.6 to 6.0 nm
FOV	14°
Pixel size	1.5x1.5 m^2

Table G.4: HySpex sensor data-sheet (nominal operations, as performed by DLR - without FOV expander).

G.5 World-View-II

World-View-II is a commercial optical satellite owned and operated by DigitalGlobe since 2009. It offers panchromatic imagery at a resolution of less than 50 centimetres and eight band multi-spectral imagery (see table G.5) [Kramer, 2014].

World-View-II (Multi-spectral)		
Bands	Violet	420 to 450 nm
	Blue	450 to 510 nm
	Green	510 to 580 nm
	Yellow	580 to 620 nm
	Red	630 to 690 nm
	Red-Edge	700 to 744 nm
	NIR 1	770 to 900 nm
	NIR 2	860 to 1040 nm
FOV	1.28°	
Pixel size	1.8x1.8 m^2 (Nadir)	
Swath	16.4 km	
Data Quantisation	11 bits	
Revisit-time	1 to 3 days	

Table G.5: Overview of the World-View-II sensor characteristics (multi-spectral camera only).

Appendix H

Vocabulary and Acronyms

This appendix presents the main definitions and notions related to quality assessment. Unless specified otherwise, they are coming from the international vocabulary of metrology (VIM) as defined by the Joint Committee for Guides in Metrology (JCGM)[1] [JCGM, 2012]. When not explicitly stated otherwise, all these definitions are relating to measurements: e.g. 'Accuracy' refers implicitly to the *accuracy of a measurement*.

Accuracy

Closeness of agreement between a measured quantity value and a true quantity value of a measurand (see figure: H.1).

Bias

Estimate of the component of measurement error that in replicate measurements remains constant or varies in a predictable manner (see figure: H.1).

Error

Measured quantity value a reference quantity value (see figure H.1).

Precision

Closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions (see figure: H.1).

Processing Chain

Set of processing steps applied to the data in order to make it usable in terms of information extraction, see section 2.3. Note that this notion is closely related to the International Vocabulary of Metrology (VIM) definition of *measuring chain* series of elements of a measuring system constituting a single path of the signal from a sensor to an output element. The remote sensing

[1]The JCGM is made up of representative of the Bureau International des poids et mesures (BIPM), the International Electrotechnical Commission (IEC), the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), International Laboratory Accreditation Cooperation (ILAC), the International Organization for Standardization (ISO), the International Union of Pure and Applied Chemistry (IUPAC), the International Union of Pure and Applied Physics (IUPAP), and the International Organization of Legal Metrology (OIML).

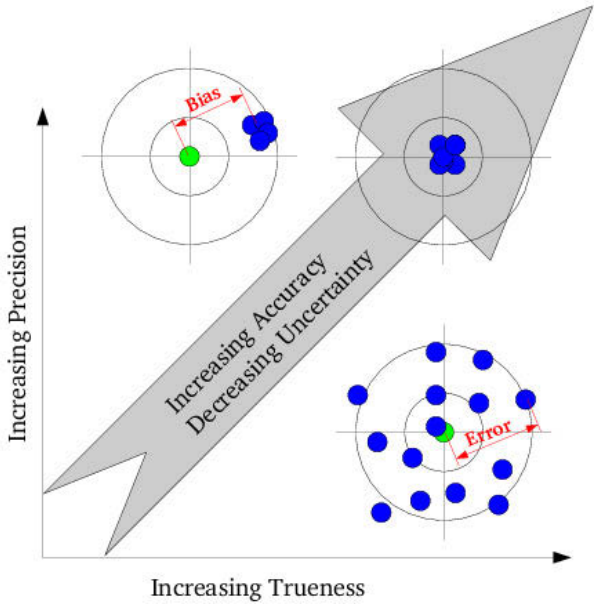

Figure H.1: Uncertainty and accuracy, as defined in Guide to the Expression of Measurement (GUM) [JCGM, 2008a].

community do however almost uniquely use the terms 'processing chain' or 'pre-processing of the data'.

Quality Indicator (QI)

Additional information providing sufficient information to allow all users to readily evaluate the 'fitness for purpose' of the data or derived product [QA4EO, 2010].

Reference Data

Data related to a property of a phenomenon, body, or substance, or to a system of components of known composition or structure, obtained from an identified source, critically evaluated, and verified for accuracy. It is not limited only to a measurement.

Repeatability

Closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement.

Reproducibility

Closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurement.

Stability of a measuring instrument

Property of a measuring instrument, whereby its metrological properties remain constant in time.

Traceability of a data-set

A documented and quantifiable assessment of evidence demonstrating the level of traceability to internationally agreed (where possible SI) reference standards [QA4EO, 2010]. This notion is related the *metrological traceability* as defined in the VIM: 'property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty.'

Trueness

Closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value (see figure: H.1).

Types of Uncertainty Evaluation Methods

The JCGM identifies two different types of uncertainty evaluation:

- Type A evaluation of measurement uncertainty evaluation of a component of measurement uncertainty by a statistical analysis of measured quantity values obtained under defined measurement conditions.
- Type B evaluation of measurement uncertainty evaluation of a component of measurement uncertainty determined by means other than a Type A evaluation of measurement uncertainty.

Uncertainty

Non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used (see figure: H.1).

Worse Case Scenario of a measurement or data processing

This denotes situations where the uncertainty is the greater. In other words, this is an estimation of the lower bound of the data quality. This is an internal definition.

Acronyms	Signification
AHS	Airborne Hyperspectral Scanner [Fernández-Renau et al., 2005]
BRDF	Bi-Directional Reflectance Function
CCD	Charged-Coupled Devices
CEOS	Committee on Earth Observation Satellites
CI	Confidence Interval
CPAM	Corresponding Pixels Automated Matcher
DLR	Deutsches Zentrum für Luft- und Raumfahrt
DCC	Dense Convective Clouds
DEM	Digital Elevation Model
DSM	Digital Surface Model
DWT	Discrete Wavelet Transform
EUFAR	European Facility for Airborne Research
FAA	Federal Aviation Authority
FFT	Fast-Fourier Transform
FLOPS	Floating Point Operation Per Second
FOV	Field of View
FWHM	Full Width at Half-Maximum
GFLOPS	10^9 Floating Point Operation Per Second
GIQE	General Image Quality Equation
GPS	Global Positioning System
GSD	Ground Sampling Distance
GUM	Guide to the expression of Uncertainty in Measurements
HyMap	Hyperspectral Mapper [Cocks et al., 1998]
HySpex	(DLR): Two NEO Cameras: VNIR-1600 & SWIR-320m-e
IMU	Inertial Measurement Unit
IS	Imaging Spectroscopy
ISU	International System Units
JCGM	Joint Committee for Guides in Metrology
LAS	Levy Alpha Stable (see section 8.4)
LiDAR	Light Detection and Ranging
MCEP	Monte Carlo Error Propagation
NCC	Normalised Cross Correlation
NDVI	Normalised Difference Vegetation Index
NIIRS	National Image Interpretability Rating Scales
NIR	Near Infra-Red
PCA	Principal Component Analysis
PDF	Probability Density Function
PICS	Pseudo Invariant Calibration Site
QA4EO	Quality Assessment for Earth Observation
QI-QL	Quality Indicator / Quality Layer
RADAR	RAdio Detection And Ranging
RBF	Radial Basis Function
RMSE	Root Mean Square Error
SA	Spectral Angle (see section 2.4.1)
SCA	Spectral Correlation Angle (see section 2.4.1)
SGA	Spectral Gradient Angle (see section 2.4.1)
SID	Spectral Information Divergence (see section 2.4.1)
SIFT	Scale Invariant Feature Transform [Lowe, 2004]
SURF	Speeded Up Robust Features [Bay et al., 2008]
SWIR	Short Wave Infra-Red
TPS	Thin Plate Splines
VIM	International Vocabulary of Metrology
VIS	Visible (part of the spectra)

Table H.1: Terms and Acronyms.

Notes on Implementations

Unless specified otherwise, all execution times relate to a $Gnu - 99$ C implementation of the algorithms. These were computed with version 4.4 of the Gnu-C-Compiler (GCC) with the optimisation tags '-O3 -march=native -m64'. Execution platform was a CentOS-6 Linux Workstation (64bits) equipped with Intel®Xeon®E5-2630 CPU at 2.3GHz.

The OpenMP API was used to perform parallelism [OMP, 2012], while the Gnu-Scientific-Library (GSL) version 1.16 was used for most scientific functions calls [Galassi and et al, 2012]). Non-linear optimisations were performed using the NLopt library version 2.4 [Johnson, 2012]. Finally, fast-Fourier transforms were computed by means of the Fastest Fourier Transform in the West (FFTW) version 3.3 [Frigo and Johnson, 2005].

Publications

A paper was submitted to the '*Computer & Geosciences*' journal and is currently under review. It summarises results from chapter 4 under the title 'A Data-Driven Approach to Quality Assessment for Hyperspectral Systems'.

A reviewed conference paper was published during the IEEE-WHISPER conference [Kerr and Fisher, 2013]. Other conference contributions on a non-full-paper review or on a non-first author basis include [Kerr et al., 2011b; Chevrel et al., 2012; Jordan et al., 2013].

Miscellaneous scientific communications include poster presentations at the 'QA4EO Workshop on Providing Harmonised Quality Information in Earth Observation Data by 2015' at the RAL Space in Harwell in October 2011: Kerr et al. [2011a]; Bachmann et al. [2011c] as well as the booklets for the demonstration sites elaborated during the EO-Miners project:

- Sokolov (Czech Republic) test site [Žibret et al., 2013].
- Witbank (South Africa) test site [Teršič et al., 2013b].
- Kazarman (Kyrgyzstan) test site [Teršič et al., 2013a].

Finally, contributions were made to a number of technical reports within the EUFAR project [Kerr et al., 2012], TET-1 satellite mission [Kerr et al., 2014] and EO-Miners [Ben-Dor et al., 2011], [Ben-Dor et al., 2014a], [Ben-Dor et al., 2014b], [Ehrler et al., 2013], [Falck et al., 2013], [Fischer et al., 2014a], [Fischer et al., 2014b], [Fischer et al., 2014c].

Acknowledgement

To start with, Prof. Ralf Reulke should be thanked for accepting to supervise this work and for providing valuable help and comments. Prof. Dr. Thomas Udelhoven and Dr. René Preusker are also warmly thanked for accepting of being part of the committee and providing valuable comments. The many friends and colleagues who also contributed to reviewing and improving this work also deserve my gratitude: Christian, Guillemette, both Martins and Anita.

The Applied Spectroscopy team from the German Aerospace Center campus of Oberpfaffenhofen as a whole should moreover be thanked for their valuable help and comments. Many thanks to the fellow PhD students, Adam, Anita and Christoph for the day to day conversations on both scientific and more profane topics. Christian Fischer finally deserves special thanks for his sheer kindness and willingness to help on any subject, ranging from kayaking up to geology and paperwork.

Many thanks to the former colleagues from Supaéro for their long-lasting friendship and support, regardless of time and distance: Adriaan, Claire & Tobtob, Enguerran & Geneviève, Norman, Angela & Papass, both Pôles, Nathalie & Quentin (as well as Abigaëlle), Ronan, Vince, Wourm and all the others. Similarly, many thanks to the Budo Herrsching and especially to Ben and Jeanine for incorporating me into the Bavarian way of life.

Last but not least, Tassilo Müller, who wrote his Bachelor thesis - Müller [2012] - on preliminary aspects of the choice of the most suitable spectral metric to be used within CPAM on the Sokolov-2010 scene should be thanked for his valuable help and commitment.

The funding of this work was made available thanks to several projects. The greatest part was funded through the 'EO-Miners' FP-7 project from the European Commission [EO-Miners, 2013]. Some of the remaining parts were funded through the European Commission integration activities of the EUFAR FP-7 project [EUFAR, 2014]. Finally the last stage of this work was funded by the TechnologieErprobungsTräger 1 (TET-1) satellite mission from the German Aerospace Center [TET-1, 2014].

Short Curriculum-Vitae

Page left blanck in digital version

Erklärung

Page left blanck in digital version